Sun Cluster System Administration Guide for Solaris OS

Sun microsystems

Contents

Preface

The *Sun Cluster System Administration Guide for Solaris OS* provides procedures for administering a Sun™ Cluster configuration on both SPARC® and x86 based systems.

Note – In this document, the term "x86" refers to the Intel 32-bit family of microprocessor chips and compatible microprocessor chips made by AMD.

This document is intended for experienced system administrators with extensive knowledge of Sun software and hardware. This document is not to be used as a planning or presales guide.

The instructions in this book assume knowledge of the Solaris™ Operating System and expertise with the volume manager software used with Sun Cluster.

Note – Sun Cluster software runs on two platforms, SPARC and x86. The information in this document pertains to both platforms unless otherwise specified in a special chapter, section, note, bulleted item, figure, table, or example.

Using UNIX Commands

This document contains information on commands specific to administering a Sun Cluster configuration. This document might not contain complete information on basic UNIX® commands and procedures.

See one or more of the following for this information:

- Online documentation for the Solaris software
- Other software documentation that you received with your system
- Solaris operating system (Solaris OS) man pages

Typographic Conventions

The following table describes the typographic conventions that are used in this book.

TABLE P–1 Typographic Conventions

Typeface	Meaning	Example
`AaBbCc123`	The names of commands, files, and directories, and onscreen computer output	Edit your `.login` file.
		Use `ls -a` to list all files.
		`machine_name% you have mail.`
`AaBbCc123`	What you type, contrasted with onscreen computer output	`machine_name% `**`su`**
		`Password:`
aabbcc123	Placeholder: replace with a real name or value	The command to remove a file is `rm` *filename*.
AaBbCc123	Book titles, new terms, and terms to be emphasized	Read Chapter 6 in the *User's Guide*.
		A *cache* is a copy that is stored locally.
		Do *not* save the file.
		Note: Some emphasized items appear bold online.

Shell Prompts in Command Examples

The following table shows the default UNIX system prompt and superuser prompt for the C shell, Bourne shell, and Korn shell.

TABLE P–2 Shell Prompts

Shell	Prompt
C shell	`machine_name%`
C shell for superuser	`machine_name#`
Bourne shell and Korn shell	`$`
Bourne shell and Korn shell for superuser	`#`

Related Documentation

Information about related Sun Cluster topics is available in the documentation that is listed in the following table. All Sun Cluster documentation is available at http://docs.sun.com.

Topic	Documentation
Overview	*Sun Cluster Overview for Solaris OS*
Concepts	*Sun Cluster Concepts Guide for Solaris OS*
Hardware installation and administration	*Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*
	Individual hardware administration guides
Software installation	*Sun Cluster Software Installation Guide for Solaris OS*
Data service installation and administration	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
	Individual data service guides
Data service development	*Sun Cluster Data Services Developer's Guide for Solaris OS*
System administration	*Sun Cluster System Administration Guide for Solaris OS*
Error messages	*Sun Cluster Error Messages Guide for Solaris OS*
Command and function references	*Sun Cluster Reference Manual for Solaris OS*

For a complete list of Sun Cluster documentation, see the release notes for your release of Sun Cluster software at http://docs.sun.com.

For a complete list of Sun Cluster documentation, see the release notes for your release of Sun Cluster at http://docs.sun.com.

Documentation, Support, and Training

The Sun web site provides information about the following additional resources:

- Documentation (http://www.sun.com/documentation/)
- Support (http://www.sun.com/support/)
- Training (http://www.sun.com/training/)

Getting Help

Contact your service provider if you have problems installing or using Sun Cluster. Provide the following information to your service provider.

- Your name and email address
- Your company name, address, and phone number
- The model and serial numbers of your systems
- The release number of the operating system, for example Solaris 9
- The release number of Sun Cluster, for example, Sun Cluster 3.2

Use the following commands to gather information on your system for your service provider:

Command	Function
prtconf -v	Displays the size of the system memory and reports information about peripheral devices
psrinfo -v	Displays information about processors
showrev -p	Reports which patches are installed
SPARC: prtdiag -v	Displays system diagnostic information
/usr/cluster/bin/clnode show-rev	Displays Sun Cluster release and package version information

Also, have available the contents of the /var/adm/messages file.

1
C H A P T E R 1

Introduction to Administering Sun Cluster

This chapter provides the following information about administering the cluster and the procedures for using Sun Cluster administration tool:

Overview of Administering Sun Cluster

Sun Cluster's highly available environment ensures that critical applications are available to end users. The system administrator's job is to make sure that Sun Cluster is stable and operational.

Familiarize yourself with the planning information in the *Sun Cluster Software Installation Guide for Solaris OS* and the *Sun Cluster Concepts Guide for Solaris OS* before beginning administration tasks. Sun Cluster administration is organized into tasks among the following manuals.

- Standard tasks, used to administer and maintain the cluster on a regular, perhaps daily basis. These tasks are described in this guide.

- Data service tasks, such as installation, configuration, and changing properties. These tasks are described in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

- Service tasks, such as adding or repairing storage or network hardware. These tasks are described in the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

For the most part, you can perform Sun Cluster administration tasks while the cluster is operational. If you need take a node out of the cluster or even shut down the node, you can so while the rest of the nodes continue cluster operations. For Solaris 10, unless otherwise indicated, Sun Cluster administrative tasks should be performed in the global zone. For those procedures that require the entire cluster to be shut down, schedule downtime outside normal working hours, to impose minimal impact on the system. If you plan to shut down the cluster or a cluster node, notify users in advance

Solaris Operating System (Solaris OS) Feature Restrictions

Do not enable or disable the following Sun Cluster services by using the Solaris 10 Service Management Facility (SMF) management interface.

Sun Cluster Service	FMRI
pnm	svc:/system/cluster/pnm:default
cl_event	svc:/system/cluster/cl_event:default
cl_eventlog	svc:/system/cluster/cl_eventlog:default
rpc_pmf	svc:/system/cluster/rpc_pmf:default
rpc_fed	svc:/system/cluster/rpc_fed:default
rgm	svc:/system/cluster/rgm:default
scdpm	svc:/system/cluster/scdpm:default
cl_ccra	svc:/system/cluster/cl_ccra:default
scsymon_srv	svc:/system/cluster/scsymon_srv:default
spm	svc:/system/cluster/spm:default
cl_svc_cluster_milestone	svc:/system/cluster/cl_svc_cluster_milestone:default
cl_svc_enable	svc:/system/cluster/cl_svc_enable:default
network-multipathing	svc:/system/cluster/network-multipathing

Administration Tools

You can perform administrative tasks on Sun Cluster by using a graphical user interface (GUI) or by using the command-line. The following section provides an overview of the GUI and command line tools.

Graphical User Interface

Sun Cluster supports GUI tools that you can use to perform various administrative tasks on your cluster. These GUI tools are Sun Cluster Manager and, if you are using Sun Cluster on a SPARC based system, Sun Management Center. See Chapter 12 for more information and for procedures about configuring Sun Cluster Manager and Sun Management Center. For specific information about how to use these tools, see the online help for each GUI.

Command-Line Interface

You can perform most Sun Cluster administration tasks interactively through the clsetup(1CL) utility. Whenever possible, administration procedures in this guide use the clsetup utility.

You can administer the following Main Menu items through the clsetup utility.

- Quorum
- Resource groups
- Data services
- Cluster interconnect
- Device groups and volumes
- Private host names
- New nodes
- Other cluster tasks

Other commands that you use to administer Sun Cluster are provided in the following list. See the man pages for more detailed information.

ccp(1M)	Starts remote console access to the cluster.
if_mpadm(1M)	Switches IP addresses from one adapter to another in an IP Network Multipathing group.
claccess(1CL)	Manages Sun Cluster access policies for adding nodes.
cldevice(1CL)	Manages Sun Cluster devices.
cldevicegroup(1CL)	Manages Sun Cluster device groups.
clinterconnect(1CL)	Manages the Sun Cluster interconnect.
clnasdevice(1CL)	Manages access to NAS devices for Sun Cluster.
clnode(1CL)	Manages Sun Cluster nodes.
clquorum(1CL)	Manages Sun Cluster quorum.
clreslogicalhostname(1CL)	Manages Sun Cluster resources for logical host names.
clresource(1CL)	Manages resources for Sun Cluster data services.
clresourcegroup(1CL)	Manages resources for Sun Cluster data services.
clresourcetype(1CL)	Manages resources for Sun Cluster data services.
clressharedaddress(1CL)	Manages Sun Cluster resources for shared addresses.
clsetup(1CL)	Configures Sun Cluster interactively.
clsnmphost(1CL)	Administers Sun Cluster SNMP hosts.
clsnmpmib(1CL)	Administers Sun Cluster SNMP MIB.
clsnmpuser(1CL)	Administers Sun Cluster SNMP users.

`cltelemetryattribute(1CL)`	Configures system resource monitoring.
`cluster(1CL)`	Manages the global configuration and the global status of Sun Cluster.
clvxvm	Initializes VERITAS Volume Manager (VxVM) on a Sun Cluster node and optionally performs root-disk encapsulation

In addition, you can use commands to administer the volume manager portion of Sun Cluster. These commands depend on the specific volume manager that your cluster users, either Solstice DiskSuite™, VERITAS Volume Manager, or Solaris Volume Manager™.

Preparing to Administer the Cluster

This section describes what to do to prepare for administering your cluster.

Documenting a Sun Cluster Hardware Configuration

Document the hardware aspects that are unique to your site as your Sun Cluster configuration is scaled. To save administration labor, refer to your hardware documentation when you change or upgrade the cluster. Labeling cables and connections between the various cluster components can also make administration easier.

Reduce the time required by a third-party service provider when servicing your cluster by keeping records of your original cluster configuration, and subsequent changes.

Using an Administrative Console

You can use either a dedicated workstation or a workstation connected through a management network as the *administrative console*, to administer the active cluster. Typically, you install and run the Cluster Control Panel (CCP) and graphical user interface (GUI) tools on the administrative console. For more information about the CCP, see "How to Log In to Sun Cluster Remotely" on page 21. For instructions on installing the Cluster Control Panel module for Sun Management Center and Sun Cluster Manager GUI tools, see the *Sun Cluster Software Installation Guide for Solaris OS*.

The administrative console is not a cluster node. The administrative console is used for remote access to the cluster nodes, either over the public network or through a network-based terminal concentrator.

If your SPARC cluster consists of a Sun Enterprise™ 10000 server, you must log in from the administrative console to the System Service Processor (SSP). Connect by using the *netcon.1M* command. The default method for netcon to connect with a Sun Enterprise 10000 domain is through the network interface. If the network is inaccessible, you can use netcon in "exclusive" mode

by setting the -f option. You can also send ~* during a normal netcon session. Either of the previous solutions gives you the option of toggling to the serial interface if the network becomes unreachable.

Sun Cluster does not require a dedicated administrative console, but using a console provides the following benefits:

- Enables centralized cluster management by grouping console and management tools on the same machine
- Provides potentially quicker problem resolution by Enterprise Services or your service provider

Backing Up the Cluster

Back up your cluster on a regular basis. Even though Sun Cluster provides a highly available environment, with mirrored copies of data on the storage devices, Sun Cluster is not a replacement for regular backups. Sun Cluster can survive multiple failures, but does not protect against user or program error, or catastrophic failure. Therefore, you must have a backup procedure in place to protect against data loss.

The following information should be included as part of your backup.

- All file system partitions
- All database data if you are running DBMS data services
- Disk partition information for all cluster disks
- The md.tab file if you are using Solstice DiskSuite/Solaris Volume Manager as your volume manager

Beginning to Administer the Cluster

Table 1–1 provides a starting point for administering your cluster.

TABLE 1–1 Sun Cluster 3.2 Administration Tools

Task	Tool	Instructions
Log in to the cluster remotely	Use the ccp command to launch the Cluster Control Panel (CCP). Then select one of the following icons: cconsole(1M)crlogin(1M), or ctelnet(1M).	"How to Log In to Sun Cluster Remotely" on page 21
Configure the cluster interactively	Start the clsetup(1CL) utility.	"How to Access the clsetup Utility" on page 21

TABLE 1–1 Sun Cluster 3.2 Administration Tools   *(Continued)*

Task	Tool	Instructions
Display Sun Cluster release number and version information	Use the clnode(1CL) show-rev --v -node *node*.	"How to Display Sun Cluster Release and Version Information" on page 22
Display installed resources, resource groups, and resource types	Use the following commands to display the resource information: ■ clresource(1CL) ■ clresourcegroup(1CL) ■ clresourcetype(1CL)	"How to Display Configured Resource Types, Resource Groups, and Resources" on page 24
Monitor cluster components graphically	Use Sun Cluster Manager.	Sun Cluster Manager online help
Administer some cluster components graphically	Use Sun Cluster Manager or the Sun Cluster module for Sun Management Center (which is available with Sun Cluster on SPARC based systems only).	Sun Cluster Manager or Sun Cluster module for Sun Management Center online help
Check the status of cluster components	Use the cluster(1CL) command with the status subcommand.	"How to Check the Status of Cluster Components" on page 25
Check the status of IP network multipathing groups on the public network	Use the clnode(1CL) status command with the -m option.	"How to Check the Status of the Public Network" on page 28
View the cluster configuration	Use the cluster(1CL) command with the show subcommand.	"How to View the Cluster Configuration" on page 29
Check global mount points	Use the sccheck(1M) command.	"How to Validate a Basic Cluster Configuration" on page 37
Look at the contents of Sun Cluster command logs	Examine the /var/cluster/logs/commandlog file.	"How to View the Contents of Sun Cluster Command Logs" on page 40
Look at Sun Cluster system messages	Examine the /var/adm/messages file.	"Viewing System Messages" in *System Administration Guide: Advanced Administration*
Monitor the status of Solstice DiskSuite	Use the metastat commands.	Solaris Volume Manager documentation
Monitor the status of Solaris Volume Manager if running Solaris 9 or Solaris 10	Use the metastat command.	*Solaris Volume Manager Administration Guide*

▼ How to Log In to Sun Cluster Remotely

The Cluster Control Panel (CCP) provides a launch pad for cconsole(1M), crlogin(1M), and ctelnet(1M) tools. All three tools start a multiple-window connection to a set of specified nodes. The multiple-window connection consists of a host window for each of the specified nodes and a common window. Input to the common window is sent to each of the host windows, enabling you to run commands simultaneously on all nodes of the cluster. See the ccp(1M) and cconsole(1M) man pages for more information.

You can also start cconsole, crlogin, or ctelnet sessions from the command line.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Verify that the following prerequisites are met before starting the CCP.

- Install the SUNWccon package on the administrative console.

- Make sure the PATH variable on the administrative console includes the Sun Cluster tools directory, /opt/SUNWcluster/bin, and /usr/cluster/bin. You can specify an alternate location for the tools directory by setting the $CLUSTER_HOME environment variable.

- Configure the clusters file, the serialports file, and the nsswitch.conf file if using a terminal concentrator. The files can be either /etc files or NIS or NIS+ databases. See clusters(4) and serialports(4) for more information.

2 Determine if you have a Sun Enterprise 10000 server platform.

- If no, proceed to Step 3.

- If yes, log in to the System Service Processor (SSP) and connect by using the netcon command. After the connection is made, type Shift~@ to unlock the console and gain write access.

3 Start the CCP launch pad.

From the administrative console, type the following command.

ccp *clustername*

The CCP launch pad is displayed.

4 To start a remote session with the cluster, click either the cconsole, crlogin, or ctelnet icon in the CCP launch pad.

▼ How to Access the clsetup Utility

The clsetup(1CL) utility enables you to interactively configure quorum, resource group, cluster transport, private host name, device group, and new node options for the cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser on any node in the cluster.

2   Start the `clsetup` **utility.**

```
# clsetup
```

The Main Menu is displayed.

3   Choose your configuration from the menu. Follow the onscreen instructions to complete a task.

See Also   See the `clsetup` online help for more information.

▼ How to Display Sun Cluster Patch Information

You do not need to be logged in as superuser to perform this procedure.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

▶   **Display the Sun Cluster patch information:**

```
% showrev -p
```

Sun Cluster update releases are identified by the main product patch number plus the update version.

Example 1–1   Displaying Sun Cluster Patch Information

The following example displays information about patch 110648-05.

```
% showrev -p | grep 110648
Patch: 110648-05 Obsoletes:  Requires:  Incompatibles:  Packages:
```

▼ How to Display Sun Cluster Release and Version Information

You do not need to be logged in as superuser to perform this procedure.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

◗ **Display Sun Cluster release and version information:**

% **clnode show-rev -v -***node*

This command displays Sun Cluster release number and version strings for all Sun Cluster packages.

Example 1–2 Displaying Sun Cluster Release and Version Information

The following example displays the cluster's release information and version information for all packages.

```
% clnode show-rev
3.2

% clnode show-rev -v
Sun Cluster 3.2 for Solaris 9 sparc

SUNWscr:        3.2.0,REV=2006.02.17.18.11
SUNWscu:        3.2.0,REV=2006.02.17.18.11
SUNWsczu:       3.2.0,REV=2006.02.17.18.11
SUNWscsck:      3.2.0,REV=2006.02.17.18.11
SUNWscnm:       3.2.0,REV=2006.02.17.18.11
SUNWscdev:      3.2.0,REV=2006.02.17.18.11
SUNWscgds:      3.2.0,REV=2006.02.17.18.11
SUNWscman:      3.2.0,REV=2005.10.18.08.42
SUNWscsal:      3.2.0,REV=2006.02.17.18.11
SUNWscsam:      3.2.0,REV=2006.02.17.18.11
SUNWscvm:       3.2.0,REV=2006.02.17.18.11
SUNWmdm:        3.2.0,REV=2006.02.17.18.11
SUNWscmasa:     3.2.0,REV=2006.02.17.18.11
SUNWscmautil:   3.2.0,REV=2006.02.17.18.11
SUNWscmautilr:  3.2.0,REV=2006.02.17.18.11
SUNWjfreechart: 3.2.0,REV=2006.02.17.18.11
SUNWscva:       3.2.0,REV=2006.02.17.18.11
SUNWscspm:      3.2.0,REV=2006.02.17.18.11
SUNWscspmu:     3.2.0,REV=2006.02.17.18.11
SUNWscspmr:     3.2.0,REV=2006.02.17.18.11
SUNWscderby:    3.2.0,REV=2006.02.17.18.11
SUNWsctelemetry: 3.2.0,REV=2006.02.17.18.11
SUNWscrsm:      3.2.0,REV=2006.02.17.18.11
SUNWcsc:        3.2.0,REV=2006.02.21.10.16
SUNWcscspm:     3.2.0,REV=2006.02.21.10.16
SUNWcscspmu:    3.2.0,REV=2006.02.21.10.16
SUNWdsc:        3.2.0,REV=2006.02.21.10.09
SUNWdscspm:     3.2.0,REV=2006.02.21.10.09
SUNWdscspmu:    3.2.0,REV=2006.02.21.10.09
SUNWesc:        3.2.0,REV=2006.02.21.10.11
SUNWescspm:     3.2.0,REV=2006.02.21.10.11
SUNWescspmu:    3.2.0,REV=2006.02.21.10.11
```

```
SUNWfsc:      3.2.0,REV=2006.02.21.10.06
SUNWfscspm:   3.2.0,REV=2006.02.21.10.06
SUNWfscspmu:  3.2.0,REV=2006.02.21.10.06
SUNWhsc:      3.2.0,REV=2006.02.21.10.20
SUNWhscspm:   3.2.0,REV=2006.02.21.10.20
SUNWhscspmu:  3.2.0,REV=2006.02.21.10.20
SUNWjsc:      3.2.0,REV=2006.02.21.10.22
SUNWjscman:   3.2.0,REV=2006.02.21.10.22
SUNWjscspm:   3.2.0,REV=2006.02.21.10.22
SUNWjscspmu:  3.2.0,REV=2006.02.21.10.22
SUNWksc:      3.2.0,REV=2006.02.21.10.14
SUNWkscspm:   3.2.0,REV=2006.02.21.10.14
SUNWkscspmu:  3.2.0,REV=2006.02.21.10.14
```

▼ How to Display Configured Resource Types, Resource Groups, and Resources

You can also accomplish this procedure by using the Sun Cluster Manager GUI. Refer to Chapter 12. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

Before You Begin   Users other than superuser require `solaris.clustser.read` RBAC authorization to use this subcommand.

◗ **Display the cluster's configured resource types, resource groups, and resources.**

   `% cluster show -t resource,resourcetype,resourcegroup`

   For information about individual resources, resource groups, and resource types, use the show subcommand with one of the following commands:

   - `resource`
   - `resource group`
   - `resourcetype`

Example 1–3   Displaying Configured Resource Types, Resource Groups, and Resources

The following example shows the resource types (RT Name), resource groups (RG Name), and resources (RS Name) configured for the cluster schost.

   `% cluster show resource,resourcetype,resourcegroup`

```
=== Registered Resource Types ===

Resource Type:                          SUNW.qfs
  RT_description:                        SAM-QFS Agent on SunCluster
  RT_version:                           3.1
  API_version:                          3
  RT_basedir:                           /opt/SUNWsamfs/sc/bin
  Single_instance:                      False
  Proxy:                                False
  Init_nodes:                           All potential masters
  Installed_nodes:                      <All>
  Failover:                             True
  Pkglist:                              <NULL>
  RT_system:                            False

=== Resource Groups and Resources ===

Resource Group:                         qfs-rg
  RG_description:                        <NULL>
  RG_mode:                               Failover
  RG_state:                              Managed
  Failback:                              False
  Nodelist:                              phys-schost-2 phys-schost-1

  --- Resources for Group qfs-rg ---

  Resource:                             qfs-res
    Type:                               SUNW.qfs
    Type_version:                       3.1
    Group:                              qfs-rg
    R_description:
    Resource_project_name:              default
    Enabled{phys-schost-2}:             True
    Enabled{phys-schost-1}:             True
    Monitored{phys-schost-2}:           True
    Monitored{phys-schost-1}:           True
```

▼ How to Check the Status of Cluster Components

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

Before You Begin   Users other than superuser require `solaris.cluster.read` RBAC authorization to use the status subcommand.

●   **Check the status of cluster components.**

        % **cluster status**

Example 1–4   Checking the Status of Cluster Components

The following example provides a sample of status information for cluster components returned by cluster cluster(1CL) status.

```
% cluster status
=== Cluster Nodes ===

--- Node Status ---

Node Name                                 Status
---------                                 ------
phys-schost-1                             Online
phys-schost-2                             Online


=== Cluster Transport Paths ===

Endpoint1                 Endpoint2                 Status
---------                 ---------                 ------
phys-schost-1:qfe1        phys-schost-4:qfe1        Path online
phys-schost-1:hme1        phys-schost-4:hme1        Path online


=== Cluster Quorum ===

--- Quorum Votes Summary ---

            Needed    Present    Possible
            ------    -------    --------
            3         3          4


--- Quorum Votes by Node ---

Node Name         Present         Possible         Status
---------         -------         --------         ------
phys-schost-1     1               1                Online
phys-schost-2     1               1                Online
```

```
      --- Quorum Votes by Device ---

      Device Name            Present      Possible        Status
      -----------            -------      --------        ------
      /dev/did/rdsk/d2s2     1           1               Online
      /dev/did/rdsk/d8s2     0           1               Offline


      === Cluster Device Groups ===

      --- Device Group Status ---

      Device Group Name      Primary          Secondary    Status
      -----------------      -------          ---------    ------
      schost-2               phys-schost-2    -            Degraded


      --- Spare, Inactive, and In Transition Nodes ---

      Device Group Name    Spare Nodes    Inactive Nodes    In Transistion Nodes
      -----------------    -----------    --------------    --------------------
      schost-2             -              -                 -


      === Cluster Resource Groups ===

      Group Name       Node Name        Suspended      Status
      ----------       ---------        ---------      ------
      test-rg          phys-schost-1    No             Offline
                       phys-schost-2    No             Online

      test-rg          phys-schost-1    No             Offline
                       phys-schost-2    No             Error--stop failed

      test-rg          phys-schost-1    No             Online
                       phys-schost-2    No             Online


      === Cluster Resources ===

      Resource Name    Node Name        Status         Message
      -------------    ---------        ------         -------
      test_1           phys-schost-1    Offline        Offline
                       phys-schost-2    Online         Online

      test_1           phys-schost-1    Offline        Offline
                       phys-schost-2    Stop failed    Faulted
```

```
test_1              phys-schost-1      Online         Online
                    phys-schost-2      Online         Online


Device Instance                 Node                    Status
---------------                 ----                    ------
/dev/did/rdsk/d2                phys-schost-1           Ok

/dev/did/rdsk/d3                phys-schost-1           Ok
                                phys-schost-2           Ok

/dev/did/rdsk/d4                phys-schost-1           Ok
                                phys-schost-2           Ok

/dev/did/rdsk/d6                phys-schost-2           Ok
```

▼ How to Check the Status of the Public Network

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

To check the status of the IP Network Multipathing groups, use the status subcommand of the clnode(1CL) command.

Before You Begin   Users other than superuser require solaris.cluster.read RBAC authorization to use this subcommand.

◗ **Check the status of cluster components.**

    % **clnode status -m**

Example 1–5   Checking the Public Network Status

The following example provides a sample of status information for cluster components returned by clnode status command.

```
% clnode status -m
--- Node IPMP Group Status ---

Node Name          Group Name    Status    Adapter    Status
---------          ----------    ------    -------    ------
phys-schost-1      test-rg       Online    qfe1       Online
phys-schost-2      test-rg       Online    qfe1       Online
```

▼ How to View the Cluster Configuration

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

Before You Begin   Users other than superuser require solaris.cluster.read RBAC authorization to use the status subcommand.

◗ **View the cluster configuration.**

```
% cluster show
```

To display more information by using the cluster command, use the verbose options. See the cluster(1CL) man page for details.

Example 1–6   Viewing the Cluster Configuration

The following example lists the cluster configuration.

```
% cluster show

=== Cluster ===

Cluster Name:                           cluster-1
  installmode:                            disabled
  heartbeat_timeout:                      10000
  heartbeat_quantum:                      1000
  private_netaddr:                        172.16.0.0
  private_netmask:                        255.255.248.0
  max_nodes:                              64
  max_privatenets:                        10
  global_fencing:                         Unknown
  Node List:                              phys-schost-1
  Node Zones:                             phys_schost-2:za

  === Host Access Control ===

  Cluster name:                           clustser-1
    Allowed hosts:                          phys-schost-1, phys-schost-2:za
    Authentication Protocol:                sys

  === Cluster Nodes ===

  Node Name:                              phys-schost-1
```

```
Node ID:                                      1
Type:                                         cluster
Enabled:                                      yes
privatehostname:                              clusternode1-priv
reboot_on_path_failure:                       disabled
globalzoneshares:                             3
defaultpsetmin:                               1
quorum_vote:                                  1
quorum_defaultvote:                           1
quorum_resv_key:                              0x43CB1E1800000001
Transport Adapter List:                       qfe3, hme0

--- Transport Adapters for phys-schost-1 ---

Transport Adapter:                            qfe3
  Adapter State:                              Enabled
  Adapter Transport Type:                     dlpi
  Adapter Property(device_name):              qfe
  Adapter Property(device_instance):          3
  Adapter Property(lazy_free):                1
  Adapter Property(dlpi_heartbeat_timeout):   10000
  Adapter Property(dlpi_heartbeat_quantum):   1000
  Adapter Property(nw_bandwidth):             80
  Adapter Property(bandwidth):                10
  Adapter Property(ip_address):               172.16.1.1
  Adapter Property(netmask):                  255.255.255.128
  Adapter Port Names:                         0
  Adapter Port State(0):                      Enabled

Transport Adapter:                            hme0
  Adapter State:                              Enabled
  Adapter Transport Type:                     dlpi
  Adapter Property(device_name):              hme
  Adapter Property(device_instance):          0
  Adapter Property(lazy_free):                0
  Adapter Property(dlpi_heartbeat_timeout):   10000
  Adapter Property(dlpi_heartbeat_quantum):   1000
  Adapter Property(nw_bandwidth):             80
  Adapter Property(bandwidth):                10
  Adapter Property(ip_address):               172.16.0.129
  Adapter Property(netmask):                  255.255.255.128
  Adapter Port Names:                         0
  Adapter Port State(0):                      Enabled

--- SNMP MIB Configuration on phys-schost-1 ---

SNMP MIB Name:                                Event
  State:                                      Disabled
```

```
    Protocol:                                     SNMPv2

  --- SNMP Host Configuration on phys-schost-1 ---

  --- SNMP User Configuration on phys-schost-1 ---

  SNMP User Name:                       foo
    Authentication Protocol:            MD5
    Default User:                       No

Node Name:                            phys-schost-2:za
  Node ID:                            2
  Type:                               cluster
  Enabled:                            yes
  privatehostname:                    clusternode2-priv
  reboot_on_path_failure:             disabled
  globalzoneshares:                   1
  defaultpsetmin:                     2
  quorum_vote:                        1
  quorum_defaultvote:                 1
  quorum_resv_key:                    0x43CB1E1800000002
  Transport Adapter List:             hme0, qfe3

  --- Transport Adapters for phys-schost-2 ---

  Transport Adapter:                      hme0
    Adapter State:                        Enabled
    Adapter Transport Type:               dlpi
    Adapter Property(device_name):        hme
    Adapter Property(device_instance):    0
    Adapter Property(lazy_free):          0
    Adapter Property(dlpi_heartbeat_timeout): 10000
    Adapter Property(dlpi_heartbeat_quantum): 1000
    Adapter Property(nw_bandwidth):       80
    Adapter Property(bandwidth):          10
    Adapter Property(ip_address):         172.16.0.130
    Adapter Property(netmask):            255.255.255.128
    Adapter Port Names:                   0
    Adapter Port State(0):                Enabled

  Transport Adapter:                      qfe3
    Adapter State:                        Enabled
    Adapter Transport Type:               dlpi
    Adapter Property(device_name):        qfe
    Adapter Property(device_instance):    3
    Adapter Property(lazy_free):          1
    Adapter Property(dlpi_heartbeat_timeout): 10000
    Adapter Property(dlpi_heartbeat_quantum): 1000
```

```
     Adapter Property(nw_bandwidth):            80
     Adapter Property(bandwidth):               10
     Adapter Property(ip_address):              172.16.1.2
     Adapter Property(netmask):                 255.255.255.128
     Adapter Port Names:                        0
     Adapter Port State(0):                     Enabled

   --- SNMP MIB Configuration on phys-schost-2 ---

   SNMP MIB Name:                             Event
     State:                                   Disabled
     Protocol:                                SNMPv2

   --- SNMP Host Configuration on phys-schost-2 ---

   --- SNMP User Configuration on phys-schost-2 ---

 === Transport Cables ===

 Transport Cable:                            phys-schost-1:qfe3,switch2@1
   Cable Endpoint1:                            phys-schost-1:qfe3
   Cable Endpoint2:                            switch2@1
   Cable State:                                Enabled

 Transport Cable:                            phys-schost-1:hme0,switch1@1
   Cable Endpoint1:                            phys-schost-1:hme0
   Cable Endpoint2:                            switch1@1
   Cable State:                                Enabled

 Transport Cable:                            phys-schost-2:hme0,switch1@2
   Cable Endpoint1:                            phys-schost-2:hme0
   Cable Endpoint2:                            switch1@2
   Cable State:                                Enabled

 Transport Cable:                            phys-schost-2:qfe3,switch2@2
   Cable Endpoint1:                            phys-schost-2:qfe3
   Cable Endpoint2:                            switch2@2
   Cable State:                                Enabled

 === Transport Switches ===

 Transport Switch:                           switch2
   Switch State:                               Enabled
   Switch Type:                                switch
   Switch Port Names:                          1 2
   Switch Port State(1):                       Enabled
   Switch Port State(2):                       Enabled
```

```
Transport Switch:                           switch1
  Switch State:                               Enabled
  Switch Type:                                switch
  Switch Port Names:                          1 2
  Switch Port State(1):                       Enabled
  Switch Port State(2):                       Enabled



=== Quorum Devices ===

Quorum Device Name:                         d3
  Enabled:                                    yes
  Votes:                                      1
  Global Name:                                /dev/did/rdsk/d3s2
  Type:                                       scsi
  Access Mode:                                scsi2
  Hosts (enabled):                            phys-schost-1, phys-schost-2

Quorum Device Name:                         qs1
  Enabled:                                    yes
  Votes:                                      1
  Global Name:                                qs1
  Type:                                       quorum_server
  Hosts (enabled):                            phys-schost-1, phys-schost-2
  Quorum Server Host:                         10.11.114.83
  Port:                                       9000



=== Device Groups ===

Device Group Name:                          testdg3
  Type:                                       SVM
  failback:                                   no
  Node List:                                  phys-schost-1, phys-schost-2
  preferenced:                                yes
  numsecondaries:                             1
  diskset name:                               testdg3

=== Registered Resource Types ===

Resource Type:                              SUNW.LogicalHostname:2
  RT_description:                             Logical Hostname Resource Type
  RT_version:                                 2
  API_version:                                2
  RT_basedir:                                 /usr/cluster/lib/rgm/rt/hafoip
  Single_instance:                            False
  Proxy:                                      False
  Init_nodes:                                 All potential masters
```

```
   Installed_nodes:                        <All>
   Failover:                               True
   Pkglist:                                SUNWscu
   RT_system:                              True

 Resource Type:                           SUNW.SharedAddress:2
   RT_description:                           HA Shared Address Resource Type
   RT_version:                               2
   API_version:                              2
   RT_basedir:                               /usr/cluster/lib/rgm/rt/hascip
   Single_instance:                          False
   Proxy:                                    False
   Init_nodes:                               <Unknown>
   Installed_nodes:                         <All>
   Failover:                                 True
   Pkglist:                                  SUNWscu
   RT_system:                                True

 Resource Type:                           SUNW.HAStoragePlus:4
   RT_description:                           HA Storage Plus
   RT_version:                               4
   API_version:                              2
   RT_basedir:                               /usr/cluster/lib/rgm/rt/hastorageplus
   Single_instance:                          False
   Proxy:                                    False
   Init_nodes:                               All potential masters
   Installed_nodes:                          <All>
   Failover:                                 False
   Pkglist:                                  SUNWscu
   RT_system:                                False

 Resource Type:                           SUNW.haderby
   RT_description:                           haderby server for Sun Cluster
   RT_version:                               1
   API_version:                              7
   RT_basedir:                               /usr/cluster/lib/rgm/rt/haderby
   Single_instance:                          False
   Proxy:                                    False
   Init_nodes:                               All potential masters
   Installed_nodes:                          <All>
   Failover:                                 False
   Pkglist:                                  SUNWscderby
   RT_system:                                False

 Resource Type:                           SUNW.sctelemetry
   RT_description:                           sctelemetry service for Sun Cluster
   RT_version:                               1
   API_version:                              7
```

```
   RT_basedir:                            /usr/cluster/lib/rgm/rt/sctelemetry
   Single_instance:                       True
   Proxy:                                 False
   Init_nodes:                            All potential masters
   Installed_nodes:                       <All>
   Failover:                              False
   Pkglist:                               SUNWsctelemetry
   RT_system:                             False

=== Resource Groups and Resources ===

Resource Group:                          HA_RG
  RG_description:                          <Null>
  RG_mode:                                 Failover
  RG_state:                                Managed
  Failback:                                False
  Nodelist:                                phys-schost-1 phys-schost-2

   --- Resources for Group HA_RG ---

   Resource:                             HA_R
     Type:                               SUNW.HAStoragePlus:4
     Type_version:                       4
     Group:                              HA_RG
     R_description:
     Resource_project_name:              SCSLM_HA_RG
     Enabled{phys-schost-1}:             True
     Enabled{phys-schost-2}:             True
     Monitored{phys-schost-1}:           True
     Monitored{phys-schost-2}:           True

Resource Group:                          cl-db-rg
  RG_description:                          <Null>
  RG_mode:                                 Failover
  RG_state:                                Managed
  Failback:                                False
  Nodelist:                                phys-schost-1 phys-schost-2

   --- Resources for Group cl-db-rg ---

   Resource:                             cl-db-rs
     Type:                               SUNW.haderby
     Type_version:                       1
     Group:                              cl-db-rg
     R_description:
     Resource_project_name:              default
     Enabled{phys-schost-1}:             True
     Enabled{phys-schost-2}:             True
```

```
      Monitored{phys-schost-1}:                  True
      Monitored{phys-schost-2}:                  True

  Resource Group:                              cl-tlmtry-rg
    RG_description:                              <Null>
    RG_mode:                                     Scalable
    RG_state:                                    Managed
    Failback:                                    False
    Nodelist:                                    phys-schost-1 phys-schost-2

    --- Resources for Group cl-tlmtry-rg ---

    Resource:                                  cl-tlmtry-rs
      Type:                                      SUNW.sctelemetry
      Type_version:                              1
      Group:                                     cl-tlmtry-rg
      R_description:
      Resource_project_name:                     default
      Enabled{phys-schost-1}:                    True
      Enabled{phys-schost-2}:                    True
      Monitored{phys-schost-1}:                  True
      Monitored{phys-schost-2}:                  True

  === DID Device Instances ===

  DID Device Name:                             /dev/did/rdsk/d1
    Full Device Path:                            phys-schost-1:/dev/rdsk/c0t2d0
    Replication:                                 none
    default_fencing:                             global

  DID Device Name:                             /dev/did/rdsk/d2
    Full Device Path:                            phys-schost-1:/dev/rdsk/c1t0d0
    Replication:                                 none
    default_fencing:                             global

  DID Device Name:                             /dev/did/rdsk/d3
    Full Device Path:                            phys-schost-2:/dev/rdsk/c2t1d0
    Full Device Path:                            phys-schost-1:/dev/rdsk/c2t1d0
    Replication:                                 none
    default_fencing:                             global

  DID Device Name:                             /dev/did/rdsk/d4
    Full Device Path:                            phys-schost-2:/dev/rdsk/c2t2d0
    Full Device Path:                            phys-schost-1:/dev/rdsk/c2t2d0
    Replication:                                 none
    default_fencing:                             global

  DID Device Name:                             /dev/did/rdsk/d5
```

```
   Full Device Path:                              phys-schost-2:/dev/rdsk/c0t2d0
   Replication:                                   none
   default_fencing:                               global

DID Device Name:                                  /dev/did/rdsk/d6
   Full Device Path:                              phys-schost-2:/dev/rdsk/c1t0d0
   Replication:                                   none
   default_fencing:                               global

=== NAS Devices ===

Nas Device:                                       nas_filer1
   Type:                                          netapp
   User ID:                                       root

Nas Device:                                       nas2
   Type:                                          netapp
   User ID:                                       llai
```

▼ How to Validate a Basic Cluster Configuration

The sccheck(1M) command runs a set of checks to validate the basic configuration that is required for a cluster to function properly. If no checks fail, sccheck returns to the shell prompt. If a check fails, sccheck produces reports in either the specified or the default output directory. If you run sccheck against more than one node, sccheck will produce a report for each node and a report for multi-node checks.

The sccheck command runs in two steps: data collection and analysis. Data collection can be time consuming, depending on the system configuration. You can run sccheckin verbose mode with the -v1 flag to print progress messages. Alternately, you can use the -v2 flag to run sccheck in highly verbose mode, which prints more detailed progress messages, especially during data collection.

Note – Run sccheck after performing an administration procedure that might result in changes to devices, volume management components, or the Sun Cluster configuration.

---

1 Become superuser on any node in the cluster.

   % **su**

2 Verify the cluster configuration.

   # **sccheck**

Example 1–7 Checking the Cluster Configuration With All Checks Passing

The following example shows sccheck being run in verbose mode against nodes phys-schost-1 and phys-schost-2 with all checks passing.

```
# sccheck -v1 -h phys-schost-1,phys-schost-2

sccheck: Requesting explorer data and node report from phys-schost-1.
sccheck: Requesting explorer data and node report from phys-schost-2.
scheck: phys-schost-1: Explorer finished.
scheck: phys-schost-1: Starting single-node checks.
scheck: phys-schost-1: Single-node checks finished.
scheck: phys-schost-2: Explorer finished.
scheck: phys-schost-2: Starting single-node checks.
scheck: phys-schost-2: Single-node checks finished.
scheck: Starting multi-node checks.
scheck: Multi-node checks finished
#
```

Example 1–8 Checking the Cluster Configuration With a Failed Check

The following example shows the node phys-schost-2 in the cluster suncluster missing the mount point /global/phys-schost-1. Reports are created in the output directory /var/cluster/scheck/myReports/.

```
# scheck -v1 -h phys-schost-1,phys-schost-2 -o /var/cluster/scheck/myReports

scheck: Requesting explorer data and node report from phys-schost-1.
scheck: Requesting explorer data and node report from phys-schost-2.
scheck: phys-schost-1: Explorer finished.
scheck: phys-schost-1: Starting single-node checks.
scheck: phys-schost-1: Single-node checks finished.
scheck: phys-schost-2: Explorer finished.
scheck: phys-schost-2: Starting single-node checks.
scheck: phys-schost-2: Single-node checks finished.
scheck: Starting multi-node checks.
scheck: Multi-node checks finished.
scheck: One or more checks failed.
scheck: The greatest severity of all check failures was 3 (HIGH).
scheck: Reports are in /var/cluster/scheck/myReports.
#
# cat /var/cluster/scheck/myReports/scheck-results.suncluster.txt
...
===================================================
= ANALYSIS DETAILS =
===================================================
------------------------------------
CHECK ID : 3065
SEVERITY : HIGH
FAILURE  : Global filesystem /etc/vfstab entries are not consistent across
all Sun Cluster 3.x nodes.
ANALYSIS : The global filesystem /etc/vfstab entries are not consistent across
all nodes in this cluster.
```

```
Analysis indicates:
FileSystem '/global/phys-schost-1' is on 'phys-schost-1' but missing from 'phys-schost-2'.
RECOMMEND: Ensure each node has the correct /etc/vfstab entry for the
filesystem(s) in question.
...
 #
```

▼ How to Check the Global Mount Points

The sccheck(1M) command includes checks that examine the /etc/vfstab file for configuration errors with the cluster file system and its global mount points.

Note – Run sccheck after making cluster configuration changes that have affected devices or volume management components.

---

1   Become superuser on any node in the cluster.

% **su**

2   Verify the cluster configuration.

sccheck

Example 1–9   Checking the Global Mount Points

The following example shows the node phys-schost-2 of the cluster suncluster missing the mount point /global/schost-1. Reports are being sent to the output directory, /var/cluster/sccheck/myReports/.

```
# sccheck -v1 -h phys-schost-1,phys-schost-2 -o /var/cluster/sccheck/myReports

sccheck: Requesting explorer data and node report from phys-schost-1.
sccheck: Requesting explorer data and node report from phys-schost-2.
sccheck: phys-schost-1: Explorer finished.
sccheck: phys-schost-1: Starting single-node checks.
sccheck: phys-schost-1: Single-node checks finished.
sccheck: phys-schost-2: Explorer finished.
sccheck: phys-schost-2: Starting single-node checks.
sccheck: phys-schost-2: Single-node checks finished.
sccheck: Starting multi-node checks.
sccheck: Multi-node checks finished.
sccheck: One or more checks failed.
sccheck: The greatest severity of all check failures was 3 (HIGH).
sccheck: Reports are in /var/cluster/sccheck/myReports.
#
# cat /var/cluster/sccheck/myReports/sccheck-results.suncluster.txt
```

```
...
====================================================
= ANALYSIS DETAILS =
====================================================
------------------------------------
CHECK ID : 3065
SEVERITY : HIGH
FAILURE  : Global filesystem /etc/vfstab entries are not consistent across
all Sun Cluster 3.x nodes.
ANALYSIS : The global filesystem /etc/vfstab entries are not consistent across
all nodes in this cluster.
Analysis indicates:
FileSystem '/global/phys-schost-1' is on 'phys-schost-1' but missing from 'phys-schost-2'.
RECOMMEND: Ensure each node has the correct /etc/vfstab entry for the
filesystem(s) in question.
...
#
```
 # cat /var/cluster/sccheck/myReports/sccheck-results.phys-schost-1.txt

```
...
====================================================
= ANALYSIS DETAILS =
====================================================
------------------------------------
CHECK ID : 1398
SEVERITY : HIGH
FAILURE  : An unsupported server is being used as a Sun Cluster 3.x node.
ANALYSIS : This server may not been qualified to be used as a Sun Cluster 3.x node.
Only servers that have been qualified with Sun Cluster 3.x are supported as
Sun Cluster 3.x nodes.
RECOMMEND: Because the list of supported servers is always being updated, check with
your Sun Microsystems representative to get the latest information on what servers
are currently supported and only use a server that is supported with Sun Cluster 3.x.
...
#
```

▼ How to View the Contents of Sun Cluster Command Logs

The /var/cluster/logs/commandlog ASCII text file contains records of selected Sun Cluster commands that are executed in a cluster. The logging of commands starts automatically when you set up the cluster and ends when you shut down the cluster. Commands are logged on all nodes that are up and booted in cluster mode.

Commands that are not logged in this file include those that display the configuration and current state of the cluster.

Commands that are logged in this file include those that configure and change the current state of the cluster:

- `claccess`
- `cldevice`
- `cldevicegroup`
- `clinterconnect`
- `clnasdevice`
- `clnode`
- `clquorum`
- `clreslogicalhostname`
- `clresource`
- `clresourcegroup`
- `clresourcetype`
- `clressharedaddress`
- `clsetup`
- `clsnmphost`
- `clsnmpmib`
- `clnsmpuser`
- `cltelemetryattribute`
- `cluster`
- `scconf`
- `scdidadm`
- `scdpm`
- `scgdevs`
- `scrgadm`
- `scsetup`
- `scshutdown`
- `scswitch`

Records in the `commandlog` file can contain the following elements:

- Date and timestamp
- Name of the host from which the command was executed
- Process ID of the command
- Login name of the user who executed the command
- Command that the user executed, including all options and operands

---

Note – Command options are quoted in the `commandlog` file so that you can readily identify them and copy, paste, and execute them in the shell.

- Exit status of the executed command

> **Note –** If a command aborts abnormally with unknown results, Sun Cluster does *not* show an exit
> status in the commandlog file.

By default, the commandlog file is regularly archived once a week. To change the archiving policies for
the commandlog file, on each node in the cluster, use the crontab command. See the crontab(1) man
page for more information.

Sun Cluster maintains up to eight previously archived commandlog files on each cluster node at any
given time. The commandlog file for the current week is named commandlog. The most recent
complete week's file is named commandlog.0. The oldest complete week's file is named
commandlog.7.

◗ **View the contents of the current week's** commandlog **file, one screen at a time.**

```
# more /var/cluster/logs/commandlog
```

Example 1–10   Viewing the Contents of Sun Cluster Command Logs

The following example shows the contents of commandlog that are displayed by the more command.

```
more -lines10 /var/cluster/logs/commandlog
11/11/2006 09:42:51 phys-schost-1 5222 root START - clsetup
11/11/2006 09:43:36 phys-schost-1 5758 root START - clrg add "app-sa-1"
11/11/2006 09:43:36 phys-schost-1 5758 root END 0
11/11/2006 09:43:36 phys-schost-1 5760 root START - clrg set -y
"RG_description=Department Shared Address RG" "app-sa-1"
11/11/2006 09:43:37 phys-schost-1 5760 root END 0
11/11/2006 09:44:15 phys-schost-1 5810 root START - clrg online "app-sa-1"
11/11/2006 09:44:15 phys-schost-1 5810 root END 0
11/11/2006 09:44:19 phys-schost-1 5222 root END -20988320
12/02/2006 14:37:21 phys-schost-1 5542 jbloggs START - clrg -c -g "app-sa-1"
-y "RG_description=Joe Bloggs Shared Address RG"
12/02/2006 14:37:22 phys-schost-1 5542 jbloggs END 0
```

2

Sun Cluster and RBAC

This chapter describes role-based access control (RBAC) in relation to Sun Cluster. Topics covered include:

Setting Up and Using RBAC With Sun Cluster

Use the following table to determine the documentation to consult about setting up and using RBAC. Specific steps that you follow to set up and use RBAC with Sun Cluster are provided later in this chapter.

Task	Instructions
Learn more about RBAC	Chapter 8, "Using Roles and Privileges (Overview)," in *System Administration Guide: Security Services*
Set up, manage elements of, and use RBAC	Chapter 9, "Using Role-Based Access Control (Tasks)," in *System Administration Guide: Security Services*
Learn more about RBAC elements and tools	Chapter 10, "Role-Based Access Control (Reference)," in *System Administration Guide: Security Services*

Sun Cluster RBAC Rights Profiles

Sun Cluster Manager and selected Sun Cluster commands and options that you issue at the command line use RBAC for authorization. Sun Cluster commands and options that require RBAC authorization will require one or more of the following authorization levels. Sun Cluster RBAC rights profiles apply to both global and non-global zones.

`solaris.cluster.read`	Authorization for list, show, and other read operations.
`solaris.cluster.admin`	Authorization to change the state of a cluster object.
`solaris.cluster.modify`	Authorization to change properties of a cluster object.

For more information on the RBAC authorization required by a Sun Cluster command, see the command man page.

RBAC rights profiles include one or more RBAC authorizations. You can assign these rights profiles to users or to roles to give them different levels of access to Sun Cluster. Sun provides the following rights profiles with Sun Cluster software.

Note – The RBAC rights profiles listed in the following table continue to support the old RBAC authorizations as defined in previous Sun Cluster releases.

Rights Profile	Includes Authorizations	Role Identity Permission

Sun Cluster Commands	None, but includes a list of Sun Cluster commands that run with `euid=0`	Execute selected Sun Cluster commands that you use to configure and manage a cluster, including the following subcommands for all of the Sun Cluster commands:
		■ list
		■ show
		■ status
		
		`scha_control(1HA)`
		
		`scha_resource_get(1HA)`
		
		`scha_resource_setstatus(1HA)`
		
		`scha_resourcegroup_get(1HA)`
		
		`scha_resourcetype_get(1HA)`

Basic Solaris User	This existing Solaris rights profile contains Solaris authorizations, as well as the following: `solaris.cluster.read`	Perform list, show, and other read operations for Sun Cluster commands, as well as access Sun Cluster Manager.
Cluster Operation	This rights profile is specific to Sun Cluster and contains the following authorizations:	

Rights Profile	Includes Authorizations	Role Identity Permission
	`solaris.cluster.read`	Perform list, show, export, status, and other read operations, as well as access Sun Cluster Manager.
	`solaris.cluster.admin`	Change the state of cluster objects.
System Administrator	This existing Solaris rights profile contains the same authorizations that the Cluster Management profile contains.	Perform the same operations that the Cluster Management role identity can perform, in addition to other system administration operations.
Cluster Management	This rights profile contains the same authorizations that the Cluster Operation profile contains, as well as the following authorization:	Perform the same operations that the Cluster Operation role identity can perform, as well as change properties of a cluster object.
	`solaris.cluster.modify`	

Creating and Assigning an RBAC Role With a Sun Cluster Management Rights Profile

Use this task to create a new RBAC role with a Sun Cluster Management Rights Profile and to assign users to this new role.

▼ How to Create a Role by Using the Administrative Roles Tool

Before You Begin To create a role, you must either assume a role that has the Primary Administrator rights profile assigned to it or run as `root` user.

1 **Start the Administrative Roles tool.**

To run the Administrative Roles tool, start the Solaris Management Console, as described in "How to Assume a Role in the Solaris Management Console" in *System Administration Guide: Security Services* in System Administration Guide: Security Services. Then, open the User Tool Collection, and click the Administrative Roles icon.

2 **Start the Add Administrative Role wizard.**

Select Add Administrative Role from the Action menu to start the Add Administrative Role wizard for configuring roles.

3 **Set up a role to which the Cluster Management rights profile is assigned.**

Use the Next and Back buttons to navigate between dialog boxes. Note that the Next button does not become active until you have filled in all required fields. The last dialog box enables you to review the entered data, at which point you can use the Back button to change entries or click Finish to save the new role. The following list summarizes the dialog box fields and buttons.

Role Name	Short name of the role.
Full Name	Long version of the name.
Description	Description of the role.
Role ID Number	UID for the role, automatically incremented.
Role Shell	The profile shells that are available to roles: Administrator's C, Administrator's Bourne, or Administrator's Korn shell.
Create a role mailing list	Makes a mailing list for users who are assigned to this role.
Available Rights / Granted Rights	Assigns or removes a role's rights profiles.
	Note that the system does not prevent you from typing multiple occurrences of the same command. The attributes that are assigned to the first occurrence of a command in a rights profile have precedence and all subsequent occurrences are ignored. Use the Up and Down arrows to change the order.
Server	Server for the home directory.
Path	Home directory path.
Add	Adds users who can assume this role. Must be in the same scope.
Delete	Deletes users who are assigned to this role.

Note – You need to place this profile first in the list of profiles that are assigned to the role.

4   Add users who need to use the Sun Cluster Manager features or Sun Cluster commands to the newly created role.

You use the useradd(1M) command to add a user account to the system. The -P option assigns a role to a user's account.

5   Click Finish when you have added the users.

6   Open a terminal window and become root.

7   Start and stop the name service cache daemon.

The new role does not take effect until the name service cache daemon is restarted. After becoming root, type the following text:

```
# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

▼ How to Create a Role From the Command Line

1 **Become superuser or assume a role that provides** `solaris.cluster.admin` **RBAC authorization.**

2 **Select a method for creating a role:**

- For roles in the local scope, use the `roleadd`(1M) command to specify a new local role and its attributes.

- Alternatively, for roles in the local scope, edit the `user_attr`(4) file to add a user with `type=role`.

  Use this method for emergencies only, as it is easy to make mistakes while you are typing.

- For roles in a name service, use the `smrole`(1M) command to specify the new role and its attributes.

  This command requires authentication by superuser or a role that is capable of creating other roles. You can apply the `smrole` to all name services. This command runs as a client of the Solaris Management Console server.

3 **Start and stop the name service cache daemon.**

New roles do not take effect until the name service cache daemon is restarted. As `root`, enter the following text:

```
# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

Example 2–1 Creating a Custom Operator Role by Using the `smrole` Command

The following sequence demonstrates how a role is created with the `smrole` command. In this example, a new version of the Operator role is created that has assigned to it the standard Operator rights profile and the Media Restore rights profile.

```
% su primaryadmin
# /usr/sadm/bin/smrole add -H myHost -- -c "Custom Operator" -n oper2 -a johnDoe \
-d /export/home/oper2 -F "Backup/Restore Operator" -p "Operator" -p "Media Restore"

Authenticating as user: primaryadmin

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password ::      <type primaryadmin password>

Loading Tool: com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost
Login to myHost as user primaryadmin was successful.
Download of com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost was successful.

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password ::      <type oper2 password>
```

```
# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

To view the newly created role (and any other roles), use `smrole` with the `list` option, as follows:

```
# /usr/sadm/bin/smrole list --
Authenticating as user: primaryadmin

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password ::      <type primaryadmin password>

Loading Tool: com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost
Login to myHost as user primaryadmin was successful.
Download of com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost was successful.
root                   0              Super-User
primaryadmin           100            Most powerful role
sysadmin               101            Performs non-security admin tasks
oper2                  102            Custom Operator
```

Modifying a User's RBAC Properties

You can modify a user's RBAC properties by using either the user accounts tool or the command line. To modify a user's RBAC properties, choose one of the following procedures.

- "How to Modify a User's RBAC Properties by Using the User Accounts Tool" on page 48
- "How to Modify a User's RBAC Properties From the Command Line" on page 49

▼ How to Modify a User's RBAC Properties by Using the User Accounts Tool

Before You Begin To modify a user's properties, you must either be running the User Tool Collection as root user or assume a role that has the primary administrator rights profile assigned to it.

1 Start the User Accounts tool.

To run the user accounts tool, you need to start the Solaris Management Console, as described in "How to Assume a Role in the Solaris Management Console" in *System Administration Guide: Security Services* in *System Administration Guide: Security Services*. Then, open the User Tool Collection, and click the user accounts icon.

After the user accounts tool starts, the icons for the existing user accounts are displayed in the view pane.

2 Click the user account icon to be changed and select Properties from the Action menu (or simply double-click the user account icon).

3 Click the appropriate tab in the dialog box for the property to be changed, as follows:

- To change the roles that are assigned to the user, click the Roles tab and move the role assignment to be changed to the appropriate column: Available Roles or Assigned Roles.
- To change the rights profiles that are assigned to the user, click the Rights tab and move it to the appropriate column: Available Rights or Assigned Rights.

Note – Avoid assigning rights profiles directly to users. The preferred approach is to require users to assume roles in order to perform privileged applications. This strategy discourages users from abusing privileges.

▼ How to Modify a User's RBAC Properties From the Command Line

1 Become superuser or assume a role that provides `solaris.cluster.modify` **RBAC authorization.**

2 Choose the appropriate command:

- To change the authorizations, roles, or rights profiles that are assigned to a user who is defined in the local scope, use the usermod(1M) command.
- Alternatively, to change the authorizations, roles, or rights profiles that are assigned to a user who is defined in the local scope, edit the `user_attr` file.

  Use this method for emergencies only, as it is easy to make a mistake while you are typing.

- To change the authorizations, roles, or rights profiles that are assigned to a user who is defined in a name service, use the smuser(1M) command.

  This command requires authentication as superuser or as a role that is capable of changing user files. You can apply smuser to all name services. smuser runs as a client of the Solaris Management Console server.

3

Shutting Down and Booting a Cluster

This chapter provides information about and procedures for shutting down and booting a cluster and individual cluster nodes. For information about booting a non-global zone, see Chapter 18, "Planning and Configuring Non-Global Zones (Tasks)," in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones* in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

- "Overview of Shutting Down and Booting a Cluster" on page 51
- "Shutting Down and Booting a Single Cluster Node" on page 58
- "Repairing a Full /var File System" on page 71

For a high-level description of the related procedures in this chapter, see Table 3–1 and Table 3–2.

Overview of Shutting Down and Booting a Cluster

The Sun Cluster cluster(1CL) shutdown command stops cluster services in an orderly fashion and cleanly shuts down the entire cluster. You can use the cluster shutdown command when moving the location of a cluster. You can also use the command to shut down the cluster if an application error causes data corruption.

Note – Use the cluster shutdown command instead of the shutdown or halt commands to ensure proper shutdown of the entire cluster. The Solaris shutdown command is used with the clnode(1CL) evacuate command to shut down individual nodes. See "How to Shut Down a Cluster" on page 52 or "Shutting Down and Booting a Single Cluster Node" on page 58 for more information.

The cluster shutdown command stops all nodes in a cluster by performing the following actions:

1. Takes offline all running resource groups.
2. Unmounts all cluster file systems.
3. Shuts down active device services.

4. Runs `init 0` and brings all nodes to the OpenBoot™ PROM `ok` prompt on a SPARC based system or to the GRUB menu on an x86 based system. The GRUB menus are described in more detail in Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration* in *System Administration Guide: Basic Administration*.

Note – If necessary, you can boot a node in noncluster mode so that the node does not participate in cluster membership. Noncluster mode is useful when installing cluster software or for performing certain administrative procedures. See "How to Boot a Cluster Node in Noncluster Mode" on page 68 for more information.

TABLE 3–1 Task List: Shutting Down and Booting a Cluster

Task	For Instructions
Stop the cluster. -Use `cluster(1CL) shutdown`	See "How to Shut Down a Cluster" on page 52
Start the cluster by booting all nodes. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	See "How to Boot a Cluster" on page 54
Reboot the cluster. - Use `cluster shutdown`. At the `Press any key to continue` message, boot each node individually by pressing a key. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	See "How to Reboot a Cluster" on page 55

▼ How to Shut Down a Cluster

Caution – Do not use `send brk` on a cluster console to shut down a cluster node. The command is not supported within a cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   SPARC: If your cluster is running Oracle Parallel Server or Oracle Real Application Clusters (RAC), shut down all instances of the database.

Refer to the Oracle Parallel Server or Oracle RAC product documentation for shutdown procedures.

2 **Become superuser or assume a role that provides** `solaris.cluster.admin` **RBAC authorization on any node in the cluster.**

3 **Shut down the cluster immediately.**

From a single node in the cluster, type the following command.

```
# cluster shutdown -g0 -y
```

4 **Verify that all nodes are showing the** `ok` **prompt on a SPARC-based system or a GRUB menu on an x86 based system.**

Do not power off any nodes until all cluster nodes are at the ok prompt on a SPARC-based system or in a Boot Subsystem on an x86 based system.

```
# cluster status -t node
```

5 **If necessary, power off the nodes.**

Example 3–1 SPARC: Shutting Down a Cluster

The following example shows the console output when normal cluster operation is stopped and all nodes are shut down so that the ok prompt is shown. The -g 0 option sets the shutdown grace period to zero, and -y provides an automatic yes response to the confirmation question. Shutdown messages also appear on the consoles of the other nodes in the cluster.

```
# cluster shutdown -g0 -y
Wed Mar 10 13:47:32 phys-schost-1 cl_runtime:
WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
The system is down.
syncing file systems... done
Program terminated
ok
```

Example 3–2 x86: Shutting Down a Cluster

The following example shows the console output when normal cluster operation is stopped all nodes are shut down. In this example, the ok prompt is not displayed on all of the nodes. The -g 0 option sets the shutdown grace period to zero, and -y provides an automatic yes response to the confirmation question. Shutdown messages also appear on the consoles of the other nodes in the cluster.

```
# cluster shutdown -g0 -y
May  2 10:32:57 phys-schost-1 cl_runtime:
WARNING: CMM: Monitoring disabled.
```

```
root@phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
failfasts already disabled on node 1
Print services already stopped.
May  2 10:33:13 phys-schost-1 syslogd: going down on signal 15
The system is down.
syncing file systems... done
Type any key to continue
```

See Also   See to restart a cluster that has been shut down.

▼ How to Boot a Cluster

This procedure explains how to start a cluster whose nodes have been shut down and are at the ok prompt on SPARC systems or at the Press any key to continue message on the GRB-based x86 systems.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Boot each node into cluster mode.

- On SPARC based systems, do the following:

  ok **boot**

- On x86 based systems, do the following:

  When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

  ```
  GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
  +-------------------------------------------------------------------------+
  | Solaris 10 /sol_10_x86                                                   |
  | Solaris failsafe                                                         |
  |                                                                         |
  +-------------------------------------------------------------------------+
  Use the ^ and v keys to select which entry is highlighted.
  Press enter to boot the selected OS, 'e' to edit the
  commands before booting, or 'c' for a command-line.
  ```

Note – Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration* in *System Administration Guide: Basic Administration*.

2   Verify that the nodes booted without error and are online.

The cluster(1CL) status command reports the nodes' status.

```
# cluster status -t node
```

Note – If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. If this problem arises, see "How to Repair a Full /var File System" on page 71.

Example 3–3   SPARC: Booting a Cluster

The following example shows the console output when node phys-schost-1 is booted into the cluster. Similar messages appear on the consoles of the other nodes in the cluster.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: Node phys-schost-1 with votecount = 1 added.
NOTICE: Node phys-schost-2 with votecount = 1 added.
NOTICE: Node phys-schost-3 with votecount = 1 added.
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-2 (incarnation # 937690106) has become reachable.
NOTICE: Node phys-schost-3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
NOTICE: node phys-schost-1 is up; new incarnation number = 937846227.
NOTICE: node phys-schost-2 is up; new incarnation number = 937690106.
NOTICE: node phys-schost-3 is up; new incarnation number = 937690290.
NOTICE: Cluster members: phys-schost-1 phys-schost-2 phys-schost-3.
...
```

▼ How to Reboot a Cluster

Run the cluster(1CL) shutdown command to shut down the cluster, then boot the cluster with the boot(1M) command on each node.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 SPARC: If your cluster is running Oracle Parallel Server or Oracle RAC, shut down all instances of the database.

Refer to the Oracle Parallel Server or Oracle RAC product documentation for shutdown procedures.

2 Become superuser or assume a role that provides `solaris.cluster.admin` **RBAC authorization on any node in the cluster.**

3 Shut down the cluster.

From a single node in the cluster, type the following command.

```
# cluster shutdown -g0 -y
```

Each node is shut down.

Note – Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

---

4 Boot each node.

The order in which the nodes are booted is irrelevant unless you make configuration changes between shutdowns. If you make configuration changes between shutdowns, start the node with the most current configuration first.

- On SPARC based systems, do the following:

  ```
  ok boot
  ```

- On x86 based systems, do the following:

  When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

  ```
  GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
  +----------------------------------------------------------------------+
  | Solaris 10 /sol_10_x86                                                |
  | Solaris failsafe                                                      |
  |                                                                      |
  +----------------------------------------------------------------------+
  Use the ^ and v keys to select which entry is highlighted.
  Press enter to boot the selected OS, 'e' to edit the
  commands before booting, or 'c' for a command-line.
  ```

Messages appear on the booted nodes' consoles as cluster components are activated.

5   **Verify that the nodes booted without error and are online.**

The scstat command reports the nodes' status.

cluster status -t node

Note – If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. If this problem arises, see "How to Repair a Full /var File System" on page 71.

Example 3–4   SPARC: Rebooting a Cluster

The following example shows the console output when normal cluster operation is stopped, all nodes are shut down to the ok prompt, and the cluster is restarted. The -g 0 option sets the grace period to zero, and -y provides an automatic yes response to the confirmation question. Shutdown messages also appear on the consoles of other nodes in the cluster.

```
# cluster shutdown -g0 -y
Wed Mar 10 13:47:32 phys-schost-1 cl_runtime:
WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
...
The system is down.
syncing file systems... done
Program terminated
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-2 (incarnation # 937690106) has become reachable.
NOTICE: Node phys-schost-3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
...
NOTICE: Cluster members: phys-schost-1 phys-schost-2 phys-schost-3.
...
NOTICE: Node phys-schost-1: joined cluster
...
The system is coming up.  Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
```

```
The system is ready.
phys-schost-1 console login:
NOTICE: Node phys-schost-1: joined cluster
...
The system is coming up.  Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
phys-schost-1 console login:
```

Shutting Down and Booting a Single Cluster Node

Note – Use the clnode(1CL) evacuate command in conjunction with the Solaris shutdown(1M) command to shut down an individual node. Use the cluster shutdown command only when shutting down an entire cluster. For information on shutting down and booting a non-global zone, see Chapter 20, "Installing, Booting, Halting, and Uninstalling Non-Global Zones (Tasks)," in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones* in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

TABLE 3–2 Task Map: Shutting Down and Booting a Cluster Node

Task	Tool	Instructions
Stop a cluster node	Use clnode(1CL) evacuate and cluster(1CL) shutdown	"How to Shut Down a Cluster Node" on page 59
Start a node The node must have a working connection to the cluster interconnect to attain cluster membership.	Use boot or b	"How to Boot a Cluster Node" on page 61
Stop and restart (reboot) a cluster node The node must have a working connection to the cluster interconnect to attain cluster membership.	Use clnode evacuate and cluster shutdown	"How to Reboot a Cluster Node" on page 63
Boot a node so that the node does not participate in cluster membership	Use clnode evacuate and shutdown, then boot -x or b -x	"How to Boot a Cluster Node in Noncluster Mode" on page 68

▼ How to Shut Down a Cluster Node

⚠️ **Caution –** Do not use `send brk` on a cluster console to shut down a cluster node. The command is not supported within a cluster.

For information about shutting down a non-global zone, see Chapter 20, "Installing, Booting, Halting, and Uninstalling Non-Global Zones (Tasks)," in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones* in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **SPARC: If your cluster is running Oracle Parallel Server or Oracle RAC, shut down all instances of the database.**

Refer to the Oracle Parallel Server or Oracle RAC product documentation for shutdown procedures.

2 **Become superuser or assume a role that provides** `solaris.cluster.admin` **RBAC authorization on the cluster node to be shut down.**

3 **Switch all resource groups, resources, and device groups from the node being shut down to other cluster members.**

On the node to be shut down, type the following command. The `clnode evacuate` command switches over all resource groups and device groups including all non-global zones from the specified node to the next preferred node.

`# clnode evacuate` *node*

node   Specifies the node from which you are switching resource groups and device groups.

4 **Shut down the cluster node.**

On the node to be shut down, type the following command.

`# shutdown -g0 -y -i 0`

5 **Verify that the cluster node is showing the** `ok` **prompt on a SPARC based system or the** `Press any key to continue` **message on the GRUB menu on an x86 based system.**

6 **If necessary, power off the node.**

Example 3–5   SPARC: Shutting Down a Cluster Node

The following example shows the console output when node phys-schost-1 is shut down. The -T0 option sets the grace period to zero, and the -y option provides an automatic yes response to the confirmation question. Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# clnode evacuate -S -h phys-schost-1
# shutdown -T0 -y
Wed Mar 10 13:47:32 phys-schost-1 cl_runtime:
WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
Notice: rgmd is being stopped.
Notice: rpc.pmfd is being stopped.
Notice: rpc.fed is being stopped.
umount: /global/.devices/node@1 busy
umount: /global/phys-schost-1 busy
The system is down.
syncing file systems... done
Program terminated
ok
```

Example 3–6   x86: Shutting Down a Cluster Node

The following example shows the console output when node phys-schost-1 is shut down. The -T0 option sets the grace period to zero, and the -y option provides an automatic yes response to the confirmation question. Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# clnode evacuate phys-schost-1
# shutdown -T0 -y
Shutdown started.    Wed Mar 10 13:47:32 PST 2004

Changing to init state 0 - please wait
Broadcast Message from root (console) on phys-schost-1 Wed Mar 10 13:47:32...
THE SYSTEM phys-schost-1 IS BEING SHUT DOWN NOW ! ! !
Log off now or risk your files being damaged

phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
failfasts disabled on node 1
Print services already stopped.
```

```
Mar 10 13:47:44 phys-schost-1 syslogd: going down on signal 15
umount: /global/.devices/node@2 busy
umount: /global/.devices/node@1 busy
The system is down.
syncing file systems... done
WARNING: CMM: Node being shut down.
Type any key to continue
```

See Also   See "How to Boot a Cluster Node" on page 61 to restart a cluster node that has been shut down.

▼ How to Boot a Cluster Node

If you intend to shut down or reboot other, active nodes in the cluster, wait until the node you are booting has reached at least the following status:

- SPARC: If you are running Solaris 9 OS, wait for the login prompt.
- If you are running Solaris 10 OS, wait for the multi-user-server milestone to come online.

Otherwise, the node will not be available to take over services from other nodes in the cluster that you shut down or reboot. For information about booting a non-global zone, see Chapter 20, "Installing, Booting, Halting, and Uninstalling Non-Global Zones (Tasks)," in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones* in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

Note – Starting a cluster node can be affected by the quorum configuration. In a two-node cluster, you must have a quorum device configured so that the total quorum count for the cluster is three. You should have one quorum count for each node and one quorum count for the quorum device. In this situation, if the first node is shut down, the second node continues to have quorum and runs as the sole cluster member. For the first node to come back in the cluster as a cluster node, the second node must be up and running. The required cluster quorum count (two) must be present.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **To start a cluster node that has been shut down, boot the node.**

- On SPARC based systems, do the following:

  ok **boot**
- On x86 based systems, do the following:

  When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-------------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                                   |
| Solaris failsafe                                                         |
|                                                                         |
+-------------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

Messages appear on the booted nodes' consoles as cluster components are activated.

Note – A cluster node must have a working connection to the cluster interconnect to attain cluster membership.

---

2 Verify that the node has booted without error, and is online.

The cluster status command reports the status of a node.

```
# cluster status -t node
```

Note – If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. If this problem arises, see "How to Repair a Full /var File System" on page 71.

Example 3–7 SPARC: Booting a Cluster Node

The following example shows the console output when node phys-schost-1 is booted into the cluster.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-1: joined cluster
...
The system is coming up.  Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
```

```
The system is ready.
phys-schost-1 console login:
```

▼ How to Reboot a Cluster Node

If you intend to shut down or reboot other, active nodes in the cluster, wait until the node you are rebooting has reached at least the following status:

- SPARC: If you are running Solaris 9 OS, wait for the login prompt.
- If you are running Solaris 10 OS, wait for the multi-user-server milestone to come online.

Otherwise, the node will not be available to take over services from other nodes in the cluster that you shut down or reboot. For information about rebooting a non-global zone, see Chapter 20, "Installing, Booting, Halting, and Uninstalling Non-Global Zones (Tasks)," in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones* in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 SPARC: If the cluster node is running Oracle Parallel Server or Oracle RAC, shut down all instances of the database.

Refer to the Oracle Parallel Server or Oracle RAC product documentation for shutdown procedures.

2 Become superuser or assume a role the provides `solaris.cluster.admin` **RBAC authorization on the cluster node to be shut down.**

3 Shut down the cluster node by using the `clnode evacuate` **and** `cluster shutdown` **commands.**

Enter the following commands on the node to be shut down. The clnode evacuate command switches over all device groups from the specified node to the next preferred node. The command also switches all resource groups from global or non-global zones on the specified node to the next-preferred global or non-global zones on other nodes.

```
# clnode evacuate  node
# shutdown -T0 -y
```

Note – Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

---

4 Verify that the node has booted without error, and is online.

```
# cluster status -t node
```

Example 3–8 SPARC: Rebooting a Cluster Node

The following example shows the console output when node phys-schost-1 is rebooted. Messages for this node, such as shutdown and startup notification, appear on the consoles of other nodes in the cluster.

```
# clnode evacuate phys-schost-1
# shutdown -T0 -y
Shutdown started.     Wed Mar 10 13:47:32 phys-schost-1 cl_runtime:

WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 6
The system is coming down.  Please wait.
System services are now being stopped.
Notice: rgmd is being stopped.
Notice: rpc.pmfd is being stopped.
Notice: rpc.fed is being stopped.
umount: /global/.devices/node@1 busy
umount: /global/phys-schost-1 busy
The system is down.
syncing file systems... done
rebooting...
Resetting ...
,,,
Sun Ultra 1 SBus (UltraSPARC 143MHz), No Keyboard
OpenBoot 3.11, 128 MB memory installed, Serial #5932401.
Ethernet address 8:8:20:99:ab:77, Host ID: 8899ab77.
...
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-1: joined cluster
...
The system is coming up.  Please wait.
The system is ready.
phys-schost-1 console login:
```

Example 3–9 x86: Rebooting a Cluster Node

The following example shows the console output when rebooting node phys-schost-1. Messages for this node, such as shutdown and startup notification, appear on the consoles of other nodes in the cluster.

```
# clnode evacuate phys-schost-1
# shutdown -T0 -y
Shutdown started.    Wed Mar 10 13:47:32 PST 2004

Changing to init state 6 - please wait
Broadcast Message from root (console) on phys-schost-1 Wed Mar 10 13:47:32...
THE SYSTEM phys-schost-1 IS BEING SHUT DOWN NOW ! ! !
Log off now or risk your files being damaged

phys-schost-1#
INIT: New run level: 6
The system is coming down.  Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
Print services already stopped.
Mar 10 13:47:44 phys-schost-1 syslogd: going down on signal 15

umount: /global/.devices/node@2 busy
umount: /global/.devices/node@1 busy
The system is down.
syncing file systems... done
WARNING: CMM: Node being shut down.
rebooting...

ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*                                        BIOS Lan-Console 2.0
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
    Press <Ctrl><A> for SCSISelect(TM) Utility!

Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
       SCSI ID: 1 SEAGATE  ST336605LC        160
```

```
        SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,   SCSI ID: 2 SUN       StorEdge 3310     160
        SCSI ID: 3 SUN       StorEdge 3310     160


AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition            Primary Boot Subsystem, vsn 2.0

                   Current Disk Partition Information

             Part#   Status    Type      Start      Length
             ===============================================
               1     Active   X86 BOOT    2428       21852
               2              SOLARIS     24280    71662420
               3              <unused>
               4              <unused>
          Please select the partition you wish to boot: *        *


Solaris DCB

                  loading /solaris/boot.bin


SunOS Secondary Boot version 3.00

                 Solaris Intel Platform Edition Booting System


Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
```

```
     Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
     Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

                    <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>    to boot with options
or      i <ENTER>                             to enter boot interpreter
or      <ENTER>                               to boot with defaults

                  <<< timeout in 5 seconds >>>

Select (b)oot or (i)nterpreter: Size: 276915 + 22156 + 150372 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version on81-feature-patch:08/30/2003 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
WARNING: CMM: Initialization for quorum device /dev/did/rdsk/d1s2 failed with
error EACCES. Will retry later.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
WARNING: CMM: Reading reservation keys from quorum device /dev/did/rdsk/d1s2
failed with error 2.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number =
1068503958.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number =
1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #3 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
```

```
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 online
NOTICE: CMM: Retry of initialization for quorum device /dev/did/rdsk/d1s2 was
successful.
WARNING: mod_installdrv: no major number for rsmrdt
ip: joining multicasts failed (18) on clprivnet0 - will use link layer
broadcasts for multicast
The system is coming up.  Please wait.
checking ufs filesystems
/dev/rdsk/c1t0d0s5: is clean.
NIS domain name is dev.eng.mycompany.com
starting rpc services: rpcbind keyserv ypbind done.
Setting netmask of e1000g2 to 192.168.255.0
Setting netmask of e1000g3 to 192.168.255.128
Setting netmask of e1000g0 to 192.168.255.128
Setting netmask of clprivnet0 to 192.168.255.0
Setting default IPv4 interface for multicast: add net 224.0/4: gateway phys-schost-1
syslog service starting.
obtaining access to all attached disks


*******************************************************************************
*
* The X-server can not be started on display :0...
*
*******************************************************************************
volume management starting.
Starting Fault Injection Server...
The system is ready.

phys-schost-1 console login:
```

▼ How to Boot a Cluster Node in Noncluster Mode

You can boot a node so that the node does not participate in the cluster membership, that is, in noncluster mode. Noncluster mode is useful when installing the cluster software or performing certain administrative procedures, such as patching a node.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role the provides** `solaris.cluster.admin` **RBAC authorization on the cluster node to be started in noncluster mode.**

2   Shut down the node by using the `clnode evacuate` **and** `cluster shutdown` **commands.**

The `clnode evacuate` command switches over all device groups from the specified node to the next preferred node. The command also switches all resource groups from global or non-global zones on the specified node to the next-preferred global or non-global zones on other nodes.

```
# clnode evacuate  node
# cluster shutdown -T0 -y
```

3   Verify that the node is showing the `ok` **prompt on a Solaris based system or the** `Press any key to continue` **message on a GRUB menu on an x86 based system.**

4   Boot the node in noncluster mode.

- On SPARC based systems, perform the following command:

  ```
  phys-schost# boot -xs
  ```

- On x86 based system, perform the following commands:

  ```
  phys-schost# shutdown -g -y -i0

  Press any key to continue
  ```

a.   In the GRUB menu, use the arrow keys to select the appropriate Solaris entry and type e **to edit its commands.**

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-------------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                                   |
| Solaris failsafe                                                         |
|                                                                         |
+-------------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration* in *System Administration Guide: Basic Administration.*

b.   In the boot parameters screen, use the arrow keys to select the kernel entry and type e **to edit the entry.**

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-------------------------------------------------------------------------+
| root (hd0,0,a)                                                           |
| kernel /platform/i86pc/multiboot                                        |
```

```
| module /platform/i86pc/boot_archive                                     |
+------------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

c. **Add** -x **to the command to specify that the system boot into noncluster mode.**

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. ]

grub edit> kernel /platform/i86pc/multiboot -x
```

d. **Press the Enter key to accept the change and return to the boot parameters screen.**

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+------------------------------------------------------------------------+
| root (hd0,0,a)                                                          |
| kernel /platform/i86pc/multiboot -x                                     |
| module /platform/i86pc/boot_archive                                     |
+------------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.-
```

e. **Type** b **to boot the node into noncluster mode.**

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -x option to the kernel boot parameter command.

Example 3–10 SPARC: Booting a Cluster Node in Noncluster Mode

The following example shows the console output when node phys-schost-1 is shut down and restarted in noncluster mode. The -g0 option sets the grace period to zero, -y provides an automatic yes response to the confirmation question, and -i0 invokes run level 0 (zero). Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# clnode evacuate phys-schost-1
# cluster shutdown -T0 -y
Shutdown started.    Wed Mar 10 13:47:32 phys-schost-1 cl_runtime:
```

```
WARNING: CMM monitoring disabled.
phys-schost-1#
...
rg_name = schost-sa-1 ...
offline node = phys-schost-2 ...
num of node = 0 ...
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
System services are now being stopped.
Print services stopped.
syslogd: going down on signal 15
...
The system is down.
syncing file systems... done
WARNING: node phys-schost-1 is being shut down.
Program terminated

ok boot -x
...
Not booting as part of cluster
...
The system is ready.
phys-schost-1 console login:
```

Repairing a Full /var **File System**

Both Solaris software and Sun Cluster software write error messages to the /var/adm/messages file, which over time can fill the /var file system. If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. Additionally, you might not be able to log in to the node.

▼ How to Repair a Full /var **File System**

If a node reports a full /var file system and continues to run Sun Cluster services, use this procedure to clear the full file system. Refer to "Viewing System Messages" in *System Administration Guide: Advanced Administration* in *System Administration Guide: Advanced Administration* for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser on the cluster node with the full /var **file system.**

2   Clear the full file system.

For example, delete nonessential files that are contained in the file system.

4

Data Replication Approaches

This chapter describes available data replication approaches with Sun Cluster. You must understand both host-based and storage-based data replication before you can select the combination of replication approaches that best serves your cluster.

This release of Sun Cluster supports the following releases of Sun's availability suite software:

- Sun StorageTek Availability Suite 4
- Sun StorEdge Availability Suite 3.2.1

In this manual, references to Sun StorageTek Availability Suite software also apply to Sun StorEdge Availability Suite software unless specifically stated otherwise.

This chapter contains the following sections:

Understanding Data Replication

Data replication is the copying of data from a primary storage device to a backup or secondary device. If the primary device fails, your data is available from the secondary device. In this way, data replication helps assure high availability and disaster tolerance for your cluster.

Sun Cluster supports the following approaches to data replication:

- *Host-based data replication* uses special software to replicate disk volumes between geographically dispersed nodes in real time. Remote mirror replication enables data from the master volume of the primary node to be replicated to the master volume of the geographically dispersed secondary node. A remote mirror bitmap tracks differences between the master volume on the primary disk and the master volume on the second disk.

  Host-based data replication is a less expensive data replication solution because it does not require storage arrays. Host-based data replication uses locally-attached disks. However, host-based data replication uses host resources to perform data replication and does not support scalable applications like Oracle RAC. For more information about using host-based data replication in a campus cluster environment, see "Using Host-Based Data Replication " on page 74. For more information about using host-based data replication between two or more clusters, see *Sun Cluster Geographic Edition Data Replication Guide for Hitachi TrueCopy*.

- *Storage-based data replication* uses special software to move the work of data replication off the cluster nodes and onto the storage device. This software relocation frees some node processing power to serve cluster requests. Storage-based data replication can be especially important in campus cluster configurations because this type of data replication supports scalable applications and offloads the hosts. Also, storage-based replication supports scalable applications such as Oracle RAC. For more information about using storage-based data replication in a campus cluster environment see "Using Storage-Based Data Replication" on page 75. For more information about using storage-based replication between two or more clusters and the Sun Cluster GeoEdition product that automates the process, see *Sun Cluster Geographic Edition Data Replication Guide for Sun StorageTek Availability SuiteSun Cluster Geographic Edition Data Replication Guide for Sun StorageTek Availability Suite*.

Sun StorageTek Availability Suite software provides a mechanism for host-based data replication across geographically distant clusters. "Example: Configuring Host-Based Data Replication With Sun StorEdge Availability Suite or Sun StorageTek Availability Suite Software" on page 78 at the end of this chapter provides a complete example of such a cluster configuration.

Using Host-Based Data Replication

This section describes host-based data replication in a two-room campus cluster. A two-room configuration with host-based data replication is defined as follows:

- Two separate rooms.
- Both rooms with one node each and disk subsystems.
- Data replicated across disk subsystems in these rooms.
- At least one disk subsystem. This is connected to both hosts, used as a quorum device, and located in one of the rooms.

Note – The examples in this section illustrate general campus cluster configurations and are not intended to indicate required or recommended configurations. For simplicity, the diagrams and explanations concentrate only on features unique to understanding campus clustering. For example, public-network Ethernet connections are not shown.

In this configuration, the system cannot recover automatically if the quorum disk is lost. Recovery requires intervention from your Sun service provider.

Single-mode fiber

Multimode fiber
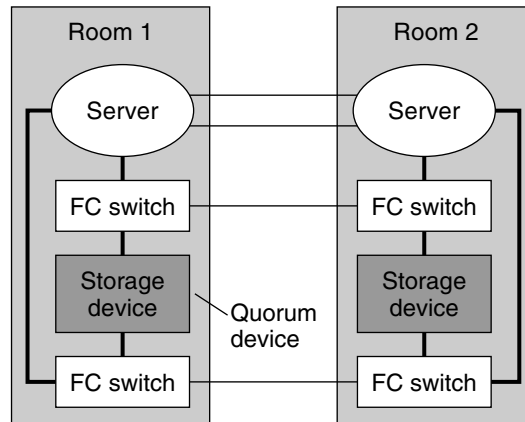
FIGURE 4–1 Two-Room Campus Cluster With Host-Based Data Replication (No Multipathing)

Figure 4–1 is similar to a standard noncampus configuration. The most obvious difference in a campus cluster is that Fibre Channel switches have been added to switch from multimode to single-mode fibers.

Using Storage-Based Data Replication

Storage-based data replication uses software installed on the storage device to manage the replication. Such software is specific to your particular storage device. Always refer to the documentation that shipped with your storage device when configuring storage-based data replication.

Depending on the software you use, you can use either automatic or manual failover with storage-based data replication. Sun Cluster supports both manual and automatic failover of the replicants with Hitachi TrueCopy software.

This section describes storage-based data replication as used in a campus cluster. Figure 4–2 shows a sample two-room configuration where data is replicated between two storage arrays. In this configuration, the primary storage array is contained in the first room, where it provides data to the nodes in both rooms. The primary storage array also provides the secondary storage array with replicated data.

During normal cluster operation, the secondary storage array is not visible to the cluster. However, if the primary storage array becomes unavailable, the secondary storage array can be manually configured into the cluster by a Sun service provider.

Note – As shown in Figure 4–2, the quorum device is on an unreplicated volume. A replicated volume cannot be used as a quorum device.
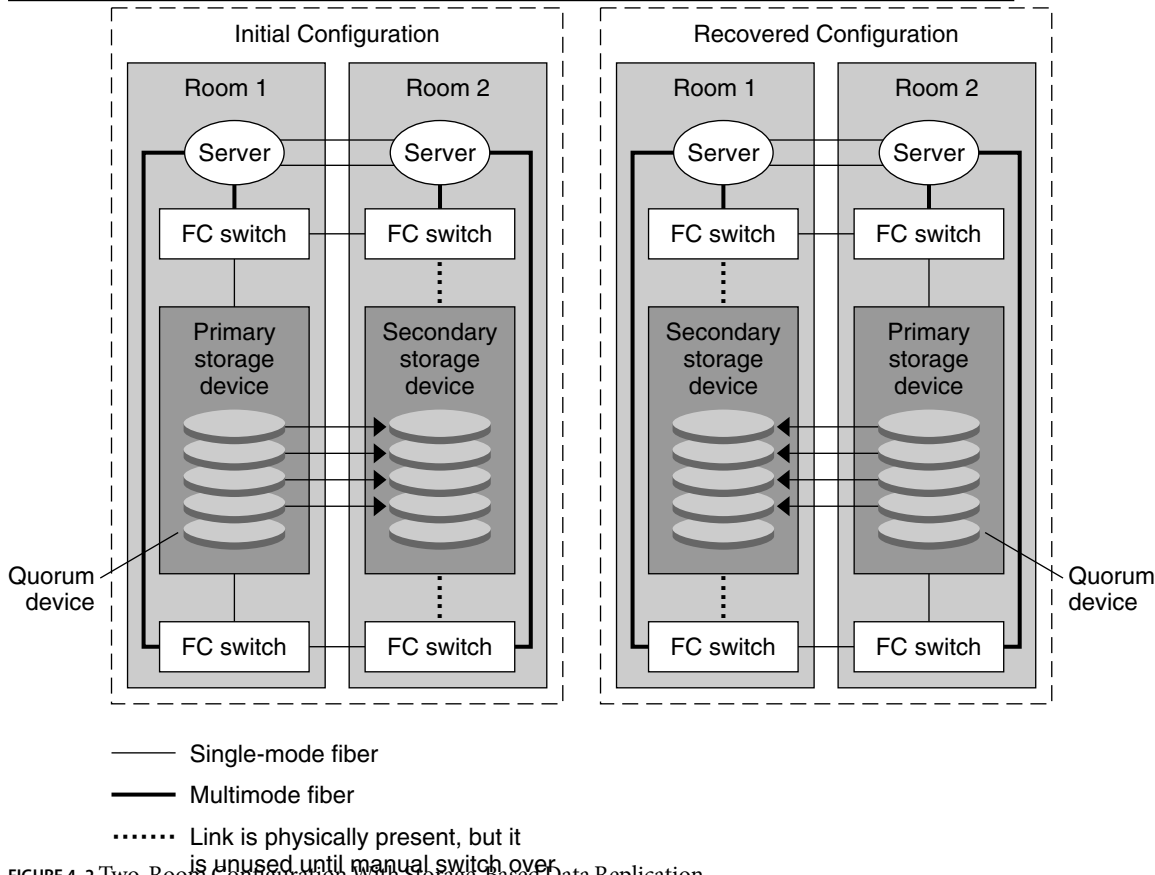
FIGURE 4–2 Two-Room Configuration With Storage-Based Data Replication

Storage-based data replication can be performed synchronously or asynchronously in the Sun Cluster environment, depending on the type of application that is used.

Requirements and Restrictions When Using Storage-Based Data Replication

To ensure data integrity, use multipathing and the proper RAID package. The following list includes considerations for implementing a campus cluster configuration that uses storage-based data replication.

- Node-to-node distance is limited by the Sun Cluster Fibre Channel and interconnect infrastructure. Contact your Sun service provider for more information about current limitations and supported technologies.

- Do not configure a replicated volume as a quorum device. Locate any quorum devices on an unreplicated volume.

- Ensure that only the primary copy of the data is visible to cluster nodes. Otherwise, the volume manager might try to access both primary and secondary copies of the data, which could result in data corruption because the secondary copy is read-only.

- When you create a disk group or diskset that is using replicated devices, use the same name for the disk group or diskset and the Hitachi TrueCopy replica pair.

- Refer to the documentation that was shipped with your storage array for information about controlling the visibility of your data copies.

- Particular application-specific data might not be suitable for asynchronous data replication. Use your understanding of your application's behavior to determine how best to replicate application-specific data across the storage devices.

- If configuring the cluster for automatic failover, use synchronous replication.

  For instructions on configuring the cluster for automatic failover of replicated volumes, see "Administering Storage-Based Replicated Devices" on page 122.

Requirements and Restrictions for Automatic Failover With Storage-Based Replication

The following restrictions apply to using storage-based data replication with automatic failover. .

- Oracle Real Application Clusters (RAC) is not supported.
- Only synchronous mode is supported.
- Replicated devices cannot be quorum devices.
- CVM and Solaris Volume Manager for Sun Cluster are not supported.

Manual Recovery Concerns When Using Storage-Based Data Replication

As with all campus clusters, those clusters that use storage-based data replication generally do not need intervention when they experience a single failure. However, if you are using manual failover and you lose the room that holds your primary storage device (as shown in Figure 4–2), problems arise in a two–node cluster. The remaining node cannot reserve the quorum device and cannot boot as a cluster member. In this situation, your cluster requires the following manual intervention:

1. Your Sun service provider must reconfigure the remaining node to boot as a cluster member.

2. You or your Sun service provider must configure an unreplicated volume of your secondary storage device as a quorum device.

3. You or your Sun service provider must configure the remaining node to use the secondary storage device as primary storage. This reconfiguration might involve rebuilding volume manager volumes, restoring data, or changing application associations with storage volumes.

Best Practices When Using TrueCopy for Storage-Based Data Replication

When setting up device groups that use the Hitachi TrueCopy software for storage-based data replication, observer the following practices:

■   Always use the highest fence level, `data`, to avoid failover to an old copy of the data.

■   Create one Hitachi TrueCopy device group per resource group. A one-to-one relationship should exist between the cluster resource group, the cluster device group, the VxVM disk group, and the Hitachi TrueCopy device group.

■   Global file-system volumes and failover file-system volumes cannot be mixed in the same Hitachi TrueCopy device group.

■   All RAID manager instances should be up and running at all times.

Example: Configuring Host-Based Data Replication With Sun StorEdge Availability Suite or Sun StorageTek Availability Suite Software

This section provides a complete example of configuring host-based data replication between clusters by using Sun StorageTek Availability Suite 3.1 or 3.2 software or Sun StorageTek Availability Suite 4.0 software. The example illustrates a complete cluster configuration for an NFS application that provides detailed information about how individual tasks can be performed. All tasks should be performed in the global zone. The example does not include all of the steps that are required by other applications or other cluster configurations.

If you use role-based access control (RBAC) instead of superuser to access the cluster nodes, ensure that you can assume an RBAC role that provides authorization for all Sun Cluster commands. This series of data replication procedures requires the following Sun Cluster RBAC authorizations if the user is not superuser:

- `solaris.cluster.modify`
- `solaris.cluster.admin`
- `solaris.cluster.read`

See Chapter 2 for more information about using RBAC roles. See the Sun Cluster man pages for the RBAC authorization that each Sun Cluster subcommand requires.

Understanding Sun StorageTek Availability Suite Software in a Cluster

This section introduces disaster tolerance and describes the data replication methods that Sun StorageTek Availability Suite software uses.

Disaster tolerance is the ability of a system to restore an application on an alternate cluster when the primary cluster fails. Disaster tolerance is based on *data replication* and *failover*. Failover is the automatic relocation of a resource group or device group from a primary cluster to a secondary cluster. If the primary cluster fails, the application and the data are immediately available on the secondary cluster.

Data Replication Methods Used by Sun StorageTek Availability Suite Software

This section describes the remote mirror replication method and the point-in-time snapshot method used by Sun StorageTek Availability Suite software. This software uses the sndradm(1RPC) and iiadm(1II) commands to replicate data.

Remote Mirror Replication

Figure 4–3 shows remote mirror replication. Data from the master volume of the primary disk is replicated to the master volume of the secondary disk through a TCP/IP connection. A remote mirror bitmap tracks differences between the master volume on the primary disk and the master volume on the secondary disk.

Primary disk   **Secondary disk**

Updated datablocks are
periodically copied to
the secondary disk

Master
volume

Master
volume

Remote mirror
bitmap volume

Remote mirror
bitmap volume

Differences between the
master volume on the
primary disk and the
master volume on the
secondary disk are
tracked on the remote
mirror bitmap volume
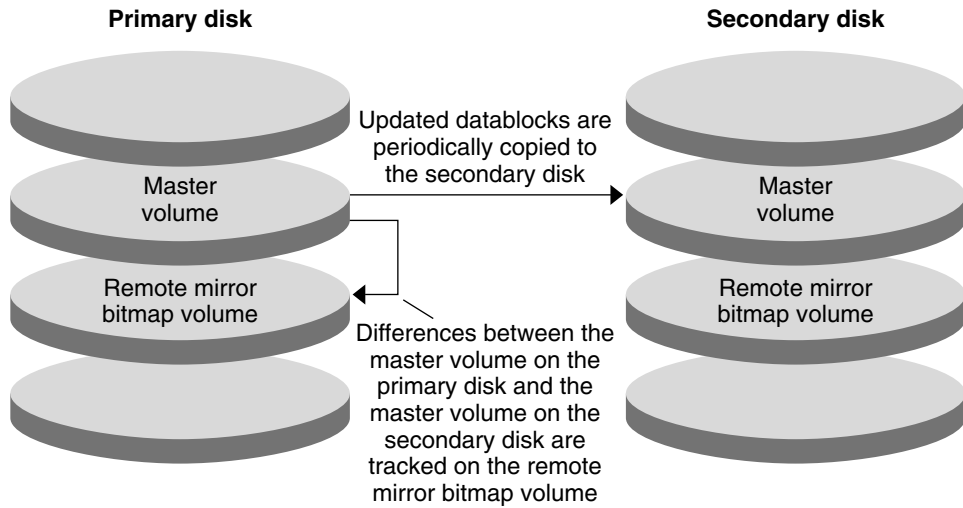
FIGURE 4–3 Remote Mirror Replication

Remote mirror replication can be performed synchronously in real time, or asynchronously. Each volume set in each cluster can be configured individually, for synchronous replication or asynchronous replication.

- In synchronous data replication, a write operation is not confirmed as complete until the remote volume has been updated.
- In asynchronous data replication, a write operation is confirmed as complete before the remote volume is updated. Asynchronous data replication provides greater flexibility over long distances and low bandwidth.

Point-in-Time Snapshot

Figure 4–4 shows point-in-time snapshot. Data from the master volume of each disk is copied to the shadow volume on the same disk. The point-in-time bitmap tracks differences between the master volume and the shadow volume. When data is copied to the shadow volume, the point-in-time bitmap is reset.
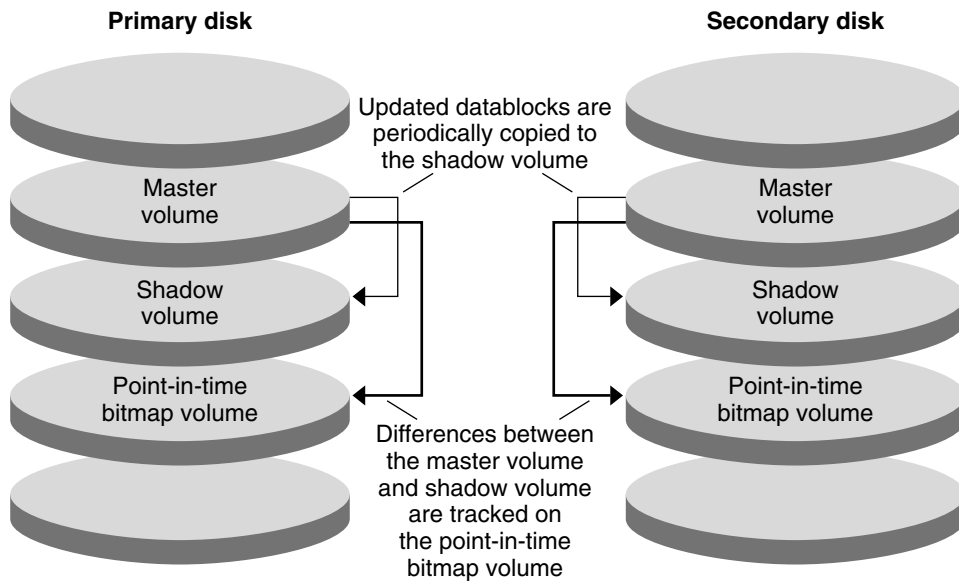
FIGURE 4–4 Point-in-Time Snapshot

Replication in the Example Configuration

Figure 4–5 illustrates how remote mirror replication and point-in-time snapshot are used in this example configuration.

Primary disk   **Secondary disk**

Master volume — Remote mirror replication → Master volume

Shadow volume ← Point-in-time replication → Shadow volume

Point-in-time bitmap volume

Remote mirror bitmap volume

Point-in-time bitmap volume

Remote mirror bitmap volume
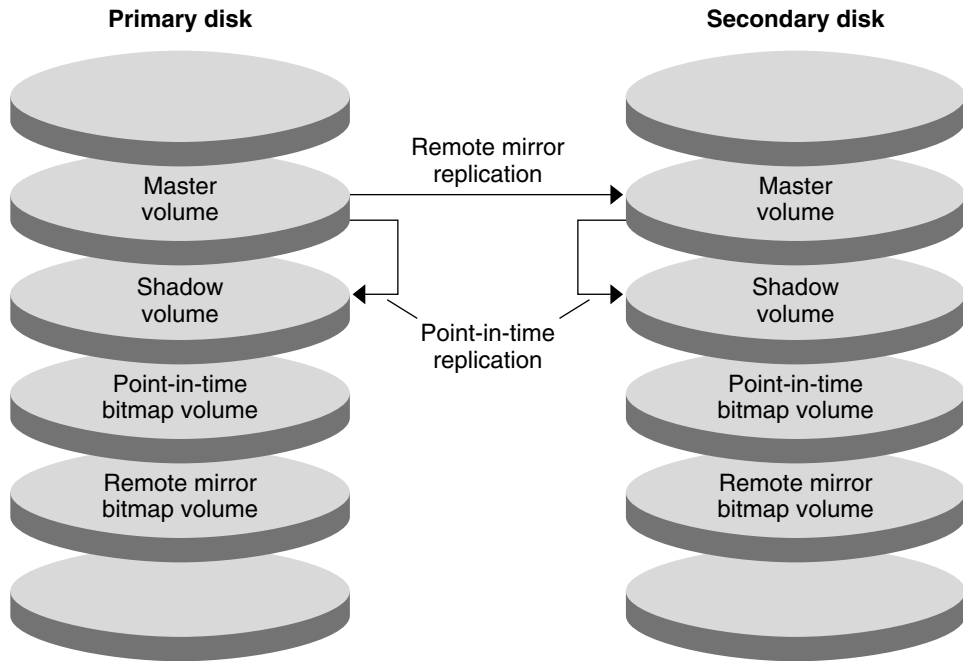
FIGURE 4–5 Replication in the Example Configuration

Guidelines for Configuring Host-Based Data Replication Between Clusters

This section provides guidelines for configuring data replication between clusters. This section also contains tips for configuring replication resource groups and application resource groups. Use these guidelines when you are configuring data replication for your cluster.

This section discusses the following topics:

- "Configuring Replication Resource Groups" on page 82
- "Configuring Application Resource Groups" on page 83
    - "Configuring Resource Groups for a Failover Application" on page 83
    - "Configuring Resource Groups for a Scalable Application" on page 85
- "Guidelines for Managing a Failover or Switchover" on page 86

Configuring Replication Resource Groups

Replication resource groups collocate the device group under Sun StorageTek Availability Suite software control with the logical host name resource. A replication resource group must have the following characteristics:

- Be a failover resource group

  A failover resource can run on only one node at a time. When a failover occurs, failover resources take part in the failover.

- Have a logical host name resource

  The logical host name must be hosted by the primary cluster. After a failover or switchover, the logical host name must be hosted by the secondary cluster. The Domain Name System (DNS) is used to associate the logical host name with a cluster.

- Have an HAStoragePlus resource

  The HAStoragePlus resource enforces the switchover of the device group when the replication resource group is switched over or failed over. Sun Cluster software also enforces the switchover of the replication resource group when the device group is switched over. In this way, the replication resource group and the device group are always colocated, or mastered by the same node.

  The following extension properties must be defined in the HAStoragePlus resource:

  - *GlobalDevicePaths*. This extension property defines the device group to which a volume belongs.

  - *AffinityOn property* = `True`. This extension property causes the device group to switch over or fail over when the replication resource group switches over or fails over. This feature is called an *affinity switchover*.

  For more information about HAStoragePlus, see the `SUNW.HAStoragePlus`(5) man page.

- Be named after the device group with which it is colocated, followed by `-stor-rg`

  For example, `devgrp-stor-rg`.

- Be online on both the primary cluster and the secondary cluster

Configuring Application Resource Groups

To be highly available, an application must be managed as a resource in an application resource group. An application resource group can be configured for a failover application or a scalable application.

Application resources and application resource groups configured on the primary cluster must also be configured on the secondary cluster. Also, the data accessed by the application resource must be replicated to the secondary cluster.

This section provides guidelines for configuring the following application resource groups:

- "Configuring Resource Groups for a Failover Application" on page 83
- "Configuring Resource Groups for a Scalable Application" on page 85

Configuring Resource Groups for a Failover Application

In a failover application, an application runs on one node at a time. If that node fails, the application fails over to another node in the same cluster. A resource group for a failover application must have the following characteristics:

- Have an HAStoragePlus resource to enforce the switchover of the device group when the application resource group is switched over or failed over

  The device group is colocated with the replication resource group and the application resource group. Therefore, the switchover of the application resource group enforces the switchover of the device group and replication resource group. The application resource group, the replication resource group, and the device group are mastered by the same node.

  Note, however, that a switchover or failover of the device group or the replication resource group does not cause a switchover or failover of the application resource group.

  - If the application data is globally mounted, the presence of an HAStoragePlus resource in the application resource group is not required but is advised.

  - If the application data is mounted locally, the presence of an HAStoragePlus resource in the application resource group is required.

    Without an HAStoragePlus resource, the switchover or failover of the application resource group would not trigger the switchover or failover of the replication resource group and device group. After a switchover or failover, the application resource group, replication resource group, and device group would not be mastered by the same node.

  For more information about HAStoragePlus, see the `SUNW.HAStoragePlus`(5) man page.

- Must be online on the primary cluster and offline on the secondary cluster

  The application resource group must be brought online on the secondary cluster when the secondary cluster takes over as the primary cluster.

Figure 4–6 illustrates the configuration of an application resource group and a replication resource group in a failover application.
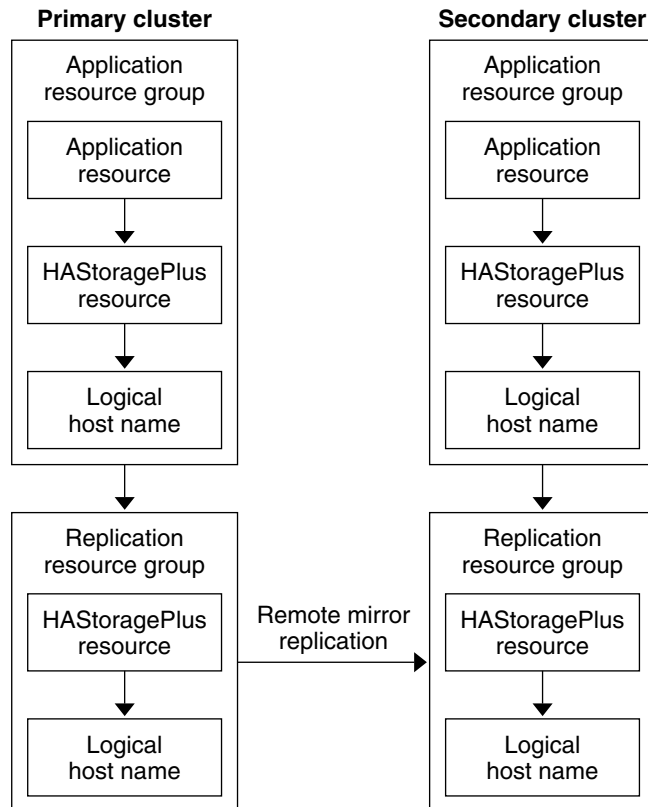
FIGURE 4–6 Configuration of Resource Groups in a Failover Application

Configuring Resource Groups for a Scalable Application

In a scalable application, an application runs on several nodes to create a single, logical service. If a node that is running a scalable application fails, failover does not occur. The application continues to run on the other nodes.

When a scalable application is managed as a resource in an application resource group, it is not necessary to collocate the application resource group with the device group. Therefore, it is not necessary to create an HAStoragePlus resource for the application resource group.

A resource group for a scalable application must have the following characteristics:

- Have a dependency on the shared address resource group

  The nodes that are running the scalable application use the shared address to distribute incoming data.

- Be online on the primary cluster and offline on the secondary cluster

Figure 4–7 illustrates the configuration of resource groups in a scalable application.
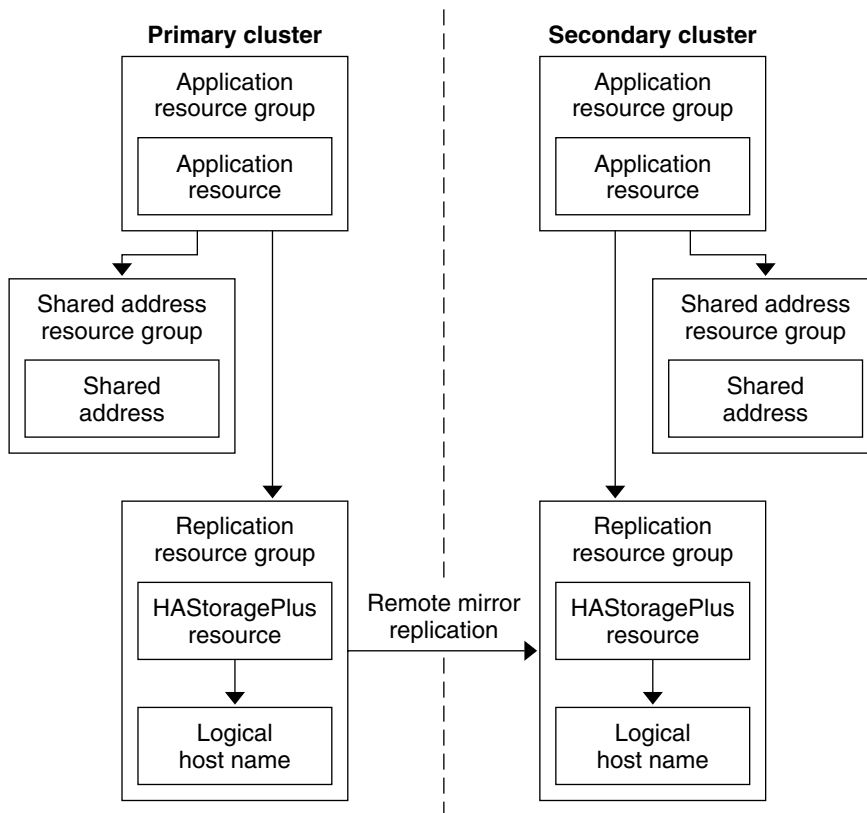
FIGURE 4–7 Configuration of Resource Groups in a Scalable Application

Guidelines for Managing a Failover or Switchover

If the primary cluster fails, the application must be switched over to the secondary cluster as soon as possible. To enable the secondary cluster to take over, the DNS must be updated.

The DNS associates a client with the logical host name of an application. After a failover or switchover, the DNS mapping to the primary cluster must be removed, and a DNS mapping to the secondary cluster must be created. Figure 4–8 shows how the DNS maps a client to a cluster.
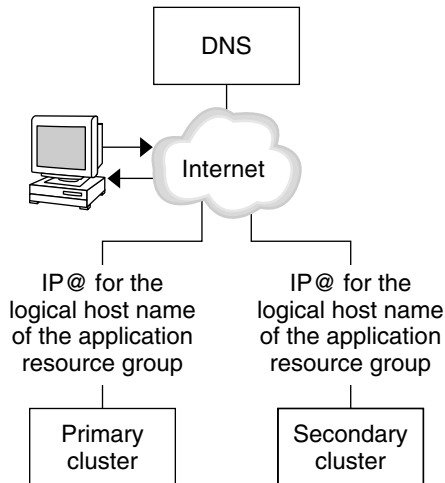
FIGURE 4–8 DNS Mapping of a Client to a Cluster

To update the DNS, use the `nsupdate` command. For information, see the `nsupdate`(1M) man page. For an example of how to manage a failover or switchover, see "Example of How to Manage a Failover or Switchover" on page 115.

After repair, the primary cluster can be brought back online. To switch back to the original primary cluster, perform the following tasks:

1. Synchronize the primary cluster with the secondary cluster to ensure that the primary volume is up-to-date.

2. Update the DNS so that clients can access the application on the primary cluster.

Task Map: Example of a Data Replication Configuration

Table 4–1 lists the tasks in this example of how data replication was configured for an NFS application by using Sun StorageTek Availability Suite software.

TABLE 4–1 Task Map: Example of a Data Replication Configuration

Task	Instructions
1. Connect and install the clusters.	"Connecting and Installing the Clusters" on page 88
2. Configure device groups, file systems for the NFS application, and resource groups on the primary cluster and on the secondary cluster.	"Example of How to Configure Device Groups and Resource Groups" on page 90

TABLE 4–1 Task Map: Example of a Data Replication Configuration   *(Continued)*

Task	Instructions
3. Enable data replication on the primary cluster and on the secondary cluster.	"How to Enable Replication on the Primary Cluster" on page 104 "How to Enable Replication on the Secondary Cluster" on page 106
4. Perform data replication.	"How to Perform a Remote Mirror Replication" on page 108 "How to Perform a Point-in-Time Snapshot" on page 110
5. Verify the data replication configuration.	"How to Verify That Replication Is Configured Correctly" on page 112

Connecting and Installing the Clusters

Figure 4–9 illustrates the cluster configuration the example configuration uses. The secondary cluster in the example configuration contains one node, but other cluster configurations can be used.
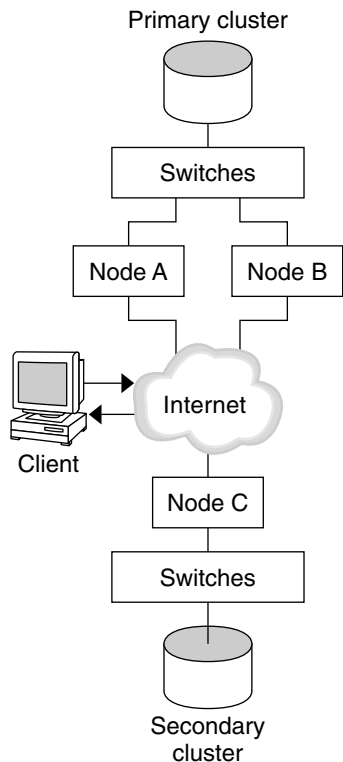
FIGURE 4–9 Example Cluster Configuration

Table 4–2 summarizes the hardware and software that the example configuration requires. The Solaris OS, Sun Cluster software, and volume manager software must be installed on the cluster nodes *before* Sun StorageTek Availability Suite software and patches are installed.

TABLE 4–2 Required Hardware and Software

Hardware or Software	Requirement
Node hardware	Sun StorageTek Availability Suite software is supported on all servers that use Solaris OS.
	For information about which hardware to use, see the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.
Disk space	Approximately 15 Mbytes.

TABLE 4–2 Required Hardware and Software   *(Continued)*

Hardware or Software	Requirement
Solaris OS	Solaris OS releases that are supported by Sun Cluster software.
	All nodes must use the same version of the Solaris OS.
	For information about installation, see the *Sun Cluster Software Installation Guide for Solaris OS*
Sun Cluster software	Sun Cluster 3.2 software.
	For information about installation, see the *Sun Cluster Software Installation Guide for Solaris OS.*
Volume manager software	Solstice DiskSuite or Solaris Volume Manager software or VERITAS Volume Manager (VxVM) software.
	All nodes must use the same version of volume manager software.
	Information about installation is in the Chapter 4, "Configuring Solaris Volume Manager Software," in *Sun Cluster Software Installation Guide for Solaris OS* and Chapter 5, "Installing and Configuring VERITAS Volume Manager," in *Sun Cluster Software Installation Guide for Solaris OS* in the *Sun Cluster Software Installation Guide for Solaris OS*
Sun StorageTek Availability Suite software	For information about how to install the software, see the installation manuals for your release of Sun StorageTek Availability Suite or Sun StorageTek Availability Suite software: ■ Sun StorEdge Availability Suite 3.1 - Sun StorEdge Availability documentation ■ Sun StorEdge Availability Suite 3.2 - Sun StorEdge Availability documentation ■ Sun StorageTek Availability Suite 4.0 – Sun StorageTek Availability documentation
Sun StorageTek Availability Suite software patches	For information about the latest patches, see `http://www.sunsolve.com`.

Example of How to Configure Device Groups and Resource Groups

This section describes how device groups and resource groups are configured for an NFS application. For additional information, see "Configuring Replication Resource Groups" on page 82 and "Configuring Application Resource Groups" on page 83.

This section contains the following procedures:

- "How to Configure a Device Group on the Primary Cluster" on page 92
- "How to Configure a Device Group on the Secondary Cluster" on page 93

The following table lists the names of the groups and resources that are created for the example configuration.

TABLE 4–3 Summary of the Groups and Resources in the Example Configuration

Group or Resource	Name	Description
Device group	`devgrp`	The device group
Replication resource group and resources	`devgrp-stor-rg`	The replication resource group
	`lhost-reprg-prim,` `lhost-reprg-sec`	The logical host names for the replication resource group on the primary cluster and the secondary cluster
	`devgrp-stor`	The HAStoragePlus resource for the replication resource group
Application resource group and resources	`nfs-rg`	The application resource group
	`lhost-nfsrg-prim,` `lhost-nfsrg-sec`	The logical host names for the application resource group on the primary cluster and the secondary cluster
	`nfs-dg-rs`	The HAStoragePlus resource for the application
	`nfs-rs`	The NFS resource

With the exception of `devgrp-stor-rg`, the names of the groups and resources are example names that can be changed as required. The replication resource group must have a name with the format *devicegroupname*`-stor-rg`.

This example configuration uses VxVM software. For information about Solstice DiskSuite or Solaris Volume Manager software, see the Chapter 4, "Configuring Solaris Volume Manager Software," in *Sun Cluster Software Installation Guide for Solaris OS* in the *Sun Cluster Software Installation Guide for Solaris OS*

The following figure illustrates the volumes that are created in the device group.
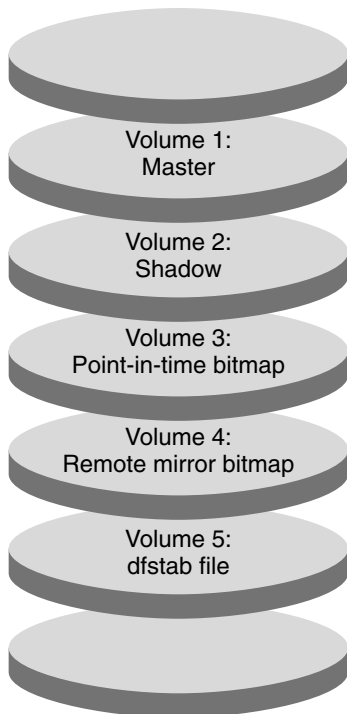
FIGURE 4–10 Volumes for the Device Group

Note – The volumes that are defined in this procedure must not include disk-label private areas, for example, cylinder 0. The VxVM software manages this constraint automatically.

▼ How to Configure a Device Group on the Primary Cluster

Before You Begin   Ensure that you have completed the following tasks:

- Read the guidelines and requirements in the following sections:
  - "Understanding Sun StorageTek Availability Suite Software in a Cluster" on page 79
  - "Guidelines for Configuring Host-Based Data Replication Between Clusters" on page 82
- Set up the primary and secondary clusters as described in "Connecting and Installing the Clusters" on page 88.

1 **Access** nodeA **as superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization.**

The node nodeA is the first node of the primary cluster. For a reminder of which node is nodeA, see Figure 4–9.

2 Create a disk group on `nodeA` **that contains volume 1,** `vol01` **through volume 4,** `vol04`.

For information about configuring a disk group by using the VxVM software, see the Chapter 5, "Installing and Configuring VERITAS Volume Manager," in *Sun Cluster Software Installation Guide for Solaris OS* in the *Sun Cluster Software Installation Guide for Solaris OSSun Cluster Software Installation Guide*.

3 Configure the disk group to create a device group.

```
nodeA# cldevicegroup create -t vxvm -n nodeA nodeB devgrp
```

The device group is called `devgrp`.

4 Create the file system for the device group.

```
nodeA# newfs /dev/vx/rdsk/devgrp/vol01 < /dev/null
nodeA# newfs /dev/vx/rdsk/devgrp/vol02 < /dev/null
```

No file system is needed for `vol03` or `vol04`, which are instead used as raw volumes.

Next Steps Go to "How to Configure a Device Group on the Secondary Cluster" on page 93.

▼ How to Configure a Device Group on the Secondary Cluster

Before You Begin Complete the procedure "How to Configure a Device Group on the Primary Cluster" on page 92.

1 Access `nodeC` **as superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization.**

2 Create a disk group on `nodeC` **that contains four volumes: volume 1,** `vol01`, **through volume 4,** `vol04`.

3 Configure the disk group to create a device group.

```
nodeC# cldevicegroup create -t vxvm -n nodeC devgrp
```

The device group is named `devgrp`.

4 Create the file system for the device group.

```
nodeC# newfs /dev/vx/rdsk/devgrp/vol01 < /dev/null
nodeC# newfs /dev/vx/rdsk/devgrp/vol02 < /dev/null
```

No file system is needed for `vol03` or `vol04`, which are instead used as raw volumes.

Next Steps Go to "How to Configure the File System on the Primary Cluster for the NFS Application" on page 93.

▼ How to Configure the File System on the Primary Cluster for the NFS Application

Before You Begin Complete the procedure "How to Configure a Device Group on the Secondary Cluster" on page 93.

1 On nodeA **and** nodeB, **become superuser or assume a role that provides** solaris.cluster.admin **RBAC authorization.**

2 On nodeA **and** nodeB, **create a mount-point directory for the NFS file system.**

For example:

```
nodeA# mkdir /global/mountpoint
```

3 On nodeA **and** nodeB, **configure the master volume to be mounted automatically on the mount point.**

Add or replace the following text in the /etc/vfstab file on nodeA and nodeB. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol01 /dev/vx/rdsk/devgrp/vol01 \
/global/mountpoint ufs 3 no global,logging
```

For a reminder of the volumes names and volume numbers that are used in the device group, see Figure 4–10.

4 On nodeA, **create a volume for the file system information that is used by the Sun Cluster HA for NFS data service.**

```
nodeA# vxassist -g devgrp make vol05 120m disk1
```

Volume 5, vol05, contains the file system information that is used by the Sun Cluster HA for NFS data service.

5 On nodeA, **resynchronize the device group with the Sun Cluster software.**

```
nodeA# cldevicegroup sync devgrp
```

6 On nodeA, **create the file system for** vol05.

```
nodeA# newfs /dev/vx/rdsk/devgrp/vol05
```

7 On nodeA **and** nodeB, **create a mount point for** vol05.

The following example creates the mount point /global/etc.

```
nodeA# mkdir /global/etc
```

8 On nodeA **and** nodeB, **configure** vol05 **to be mounted automatically on the mount point.**

Add or replace the following text in the /etc/vfstab file on nodeA and nodeB. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol05 /dev/vx/rdsk/devgrp/vol05 \
/global/etc ufs 3 yes global,logging
```

9 Mount vol05 **on** nodeA.

```
nodeA# mount /global/etc
```

10 Make vol05 **accessible to remote systems.**

    a. Create a directory called /global/etc/SUNW.nfs **on** nodeA**.**

    nodeA# **mkdir -p /global/etc/SUNW.nfs**

    b. Create the file /global/etc/SUNW.nfs/dfstab.nfs-rs **on** nodeA**.**

    nodeA# **touch /global/etc/SUNW.nfs/dfstab.nfs-rs**

    c. Add the following line to the /global/etc/SUNW.nfs/dfstab.nfs-rs **file on** nodeA**.**

    share -F nfs -o rw -d "HA NFS" /global/*mountpoint*

Next Steps   Go to "How to Configure the File System on the Secondary Cluster for the NFS Application" on page 95.

▼ How to Configure the File System on the Secondary Cluster for the NFS Application

Before You Begin   Complete the procedure "How to Configure the File System on the Primary Cluster for the NFS Application" on page 93.

1 On nodeC**, become superuser or assume a role that provides** solaris.cluster.admin **RBAC authorization.**

2 On nodeC**, create a mount-point directory for the NFS file system.**

For example:

nodeC# **mkdir /global/***mountpoint*

3 On nodeC**, configure the master volume to be mounted automatically on the mount point.**

Add or replace the following text in the /etc/vfstab file on nodeC. The text must be on a single line.

/dev/vx/dsk/devgrp/vol01 /dev/vx/rdsk/devgrp/vol01 \
/global/*mountpoint* ufs 3 no global,logging

4 On nodeC**, create a volume for the file system information that is used by the Sun Cluster HA for NFS data service.**

nodeC# **vxassist -g devgrp make vol05 120m** *disk1*

Volume 5, vol05, contains the file system information that is used by the Sun Cluster HA for NFS data service.

5 On nodeC**, resynchronize the device group with the Sun Cluster software.**

nodeC# **cldevicegroup sync devgrp**

6 On nodeC**, create the file system for** vol05**.**

nodeC# **newfs /dev/vx/rdsk/devgrp/vol05**

7   On nodeC**, create a mount point for** vol05**.**

The following example creates the mount point /global/etc.

nodeC# **mkdir /global/etc**

8   On nodeC**, configure** vol05 **to be mounted automatically on the mount point.**

Add or replace the following text in the /etc/vfstab file on nodeC. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol05 /dev/vx/rdsk/devgrp/vol05 \
/global/etc ufs 3 yes global,logging
```

9   Mount vol05 **on** nodeC**.**

nodeC# **mount /global/etc**

10   Make vol05 **accessible to remote systems.**

a.   Create a directory called /global/etc/SUNW.nfs **on** nodeC**.**

nodeC# **mkdir -p /global/etc/SUNW.nfs**

b.   Create the file /global/etc/SUNW.nfs/dfstab.nfs-rs **on** nodeC**.**

nodeC# **touch /global/etc/SUNW.nfs/dfstab.nfs-rs**

c.   Add the following line to the /global/etc/SUNW.nfs/dfstab.nfs-rs **file on** nodeC**:**

share -F nfs -o rw -d "HA NFS" /global/*mountpoint*

Next Steps   Go to "How to Create a Replication Resource Group on the Primary Cluster" on page 96.

▼   How to Create a Replication Resource Group on the Primary Cluster

Before You Begin   Complete the procedure "How to Configure the File System on the Secondary Cluster for the NFS Application" on page 95.

1   Access nodeA **as superuser or assume a role that provides** solaris.cluster.modify**,** solaris.cluster.admin**, and** solaris.cluster.read **RBAC authorization.**

2   Register the SUNW.HAStoragePlus **resource type.**

nodeA# **clresourcetype register SUNW.HAStoragePlus**

3   Create a replication resource group for the device group.

nodeA# **clresourcegroup create -n nodeA,nodeB devgrp-stor-rg**

-n nodeA,nodeB	Specifies that cluster nodes nodeA and nodeB can master the replication resource group.
devgrp-stor-rg	The name of the replication resource group. In this name, devgrp specifies the name of the device group.

4 Add a SUNW.HAStoragePlus **resource to the replication resource group.**

```
nodeA# clresource create -g devgrp-stor-rg -t SUNW.HAStoragePlus \
-p GlobalDevicePaths=devgrp \
-p AffinityOn=True \
devgrp-stor
```

-g	Specifies the resource group to which resource is added.
-p GlobalDevicePaths=	Specifies the extension property that Sun StorageTek Availability Suite software relies on.
-p AffinityOn=True	Specifies that the SUNW.HAStoragePlus resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the replication resource group fails over or is switched over, the associated device group is switched over.

For more information about these extension properties, see the SUNW.HAStoragePlus(5) man page.

5 Add a logical host name resource to the replication resource group.

```
nodeA# clreslogicalhostname create -g devgrp-stor-rg lhost-reprg-prim
```

The logical host name for the replication resource group on the primary cluster is named lhost-reprg-prim.

6 Enable the resources, manage the resource group, and bring the resource group online.

```
nodeA# clresourcegroup online -e -M -n nodeA devgrp-stor-rg
```

-e	Enables associated resources.
-M	Manages the resource group.
-n	Specifies the node on which to bring the resource group online.

7 Verify that the resource group is online.

```
nodeA# clresourcegroup status devgrp-stor-rg
```

Examine the resource group state field to confirm that the replication resource group is online on nodeA.

Next Steps Go to

▼ How to Create a Replication Resource Group on the Secondary Cluster

Before You Begin Complete the procedure

1   Access nodeC **as superuser or assume a role that provides** solaris.cluster.modify, solaris.cluster.admin, **and** solaris.cluster.read **RBAC authorization.**

2   Register SUNW.HAStoragePlus **as a resource type.**

```
nodeC# clresourcetype register SUNW.HAStoragePlus
```

3   Create a replication resource group for the device group.

```
nodeC# clresourcegroup create -n nodeC devgrp-stor-rg
```

create	Creates the resource group.
-n	Specifies the node list for the resource group.
devgrp	The name of the device group.
devgrp-stor-rg	The name of the replication resource group.

4   Add a SUNW.HAStoragePlus **resource to the replication resource group.**

```
nodeC# clresource create \
-t SUNW.HAStoragePlus \
-p GlobalDevicePaths=devgrp \
-p AffinityOn=True \
devgrp-stor
```

create	Creates the resource.
-t	Specifies the resource type.
-p GlobalDevicePaths=	Specifies the extension property that Sun StorageTek Availability Suite software relies on.
-p AffinityOn=True	Specifies that the SUNW.HAStoragePlus resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the replication resource group fails over or is switched over, the associated device group is switched over.
devgrp-stor	The HAStoragePlus resource for the replication resource group.

For more information about these extension properties, see the SUNW.HAStoragePlus(5) man page.

5   Add a logical host name resource to the replication resource group.

```
nodeC# clreslogicalhostname create -g devgrp-stor-rg lhost-reprg-sec
```

The logical host name for the replication resource group on the primary cluster is named lhost-reprg-sec.

6   Enable the resources, manage the resource group, and bring the resource group online.

```
nodeC# clresourcegroup online -e -M -n nodeC devgrp-stor-rg
```

online   Brings online.

-e   Enables associated resources.

-M   Manages the resource group.

-n   Specifies the node on which to bring the resource group online.

7 Verify that the resource group is online.

nodeC# **clresourcegroup status devgrp-stor-rg**

Examine the resource group state field to confirm that the replication resource group is online on nodeC.

Next Steps   Go to "How to Create an NFS Application Resource Group on the Primary Cluster" on page 99.

▼ How to Create an NFS Application Resource Group on the Primary Cluster

This procedure describes how application resource groups are created for NFS. This procedure is specific to this application and cannot be used for another type of application.

Before You Begin   Complete the procedure "How to Create a Replication Resource Group on the Secondary Cluster" on page 97.

1 Access nodeA **as superuser or assume a role that provides** solaris.cluster.modify, solaris.cluster.admin, **and** solaris.cluster.read **RBAC authorization.**

2 Register SUNW.nfs **as a resource type.**

nodeA# **clresourcetype register SUNW.nfs**

3 If SUNW.HAStoragePlus **has not been registered as a resource type, register it.**

nodeA# **clresourcetype register SUNW.HAStoragePlus**

4 Create an application resource group for the device group devgrp.

nodeA# **clresourcegroup create **
**-p Pathprefix=/global/etc **
**-p Auto_start_on_new_cluster=False **
**-p RG_dependencies=devgrp-stor-rg **
nfs-rg

Pathprefix=/global/etc
   Specifies the directory into which the resources in the group can write administrative files.

Auto_start_on_new_cluster=False
   Specifies that the application resource group is not started automatically.

```
RG_dependencies=devgrp-stor-rg
```
Specifies the resource group that the application resource group depends on. In this example, the application resource group depends on the replication resource group `devgrp-stor-rg`.

If the application resource group is switched over to a new primary node, the replication resource group is automatically switched over. However, if the replication resource group is switched over to a new primary node, the application resource group must be manually switched over.

```
nfs-rg
```
The name of the application resource group.

5   **Add a** `SUNW.HAStoragePlus` **resource to the application resource group.**

```
nodeA# clresource create -g nfs-rg \
-t SUNW.HAStoragePlus \
-p FileSystemMountPoints=/global/mountpoint \
-p AffinityOn=True \
nfs-dg-rs
```

```
create
```
Creates the resource.

```
-g
```
Specifies the resource group to which the resource is added.

```
-t SUNW.HAStoragePlus
```
Specifies that the resource is of the type `SUNW.HAStoragePlus`.

```
-p FileSystemMountPoints=/global/
```
Specifies that the mount point for the file system is global.

```
-p AffinityOn=True
```
Specifies that the application resource must perform an affinity switchover for the global devices and cluster file systems defined by `-p GlobalDevicePaths=`. Therefore, when the application resource group fails over or is switched over, the associated device group is switched over.

```
nfs-dg-rs
```
The name of the HAStoragePlus resource for the NFS application.

For more information about these extension properties, see the `SUNW.HAStoragePlus`(5) man page.

6   **Add a logical host name resource to the application resource group.**

```
nodeA# clreslogicalhostname create -g nfs-rg \
lhost-nfsrg-prim
```

The logical host name of the application resource group on the primary cluster is named `lhost-nfsrg-prim`.

7   **Enable the resources, manage the application resource group, and bring the application resource group online.**

   a.   **Enable the HAStoragePlus resource for the NFS application.**

   nodeA# **`clresource enable nfs-rs`**

   b.   **Bring the application resource group online on** nodeA.

   nodeA# **`clresourcegroup online -e -M -n nodeA nfs-rg`**

   online   Brings the resource group online.

   -e   Enables the associated resources.

   -M   Manages the resource group.

   -n   Specifies the node on which to bring the resource group online.

   nfs-rg   The name of the resource group.

8   **Verify that the application resource group is online.**

   nodeA# **`clresourcegroup status`**

   Examine the resource group state field to determine whether the application resource group is online for nodeA and nodeB.

Next Steps   Go to "How to Create an NFS Application Resource Group on the Secondary Cluster" on page 101.

▼ How to Create an NFS Application Resource Group on the Secondary Cluster

Before You Begin   Complete the procedure "How to Create an NFS Application Resource Group on the Primary Cluster" on page 99.

1   **Access** nodeC **as superuser or assume a role that provides** solaris.cluster.modify, solaris.cluster.admin, **and** solaris.cluster.read **RBAC authorization.**

2   **Register** SUNW.nfs **as a resource type.**

   nodeC# **`clresourcetype register SUNW.nfs`**

3   **If** SUNW.HAStoragePlus **has not been registered as a resource type, register it.**

   nodeC# **`clresourcetype register SUNW.HAStoragePlus`**

4   **Create an application resource group for the device group.**

   ```
   nodeC# clresourcegroup create \
   -p Pathprefix=/global/etc \
   -p Auto_start_on_new_cluster=False \
   ```

```
-p RG_dependencies=devgrp-stor-rg \
nfs-rg
```

create
: Creates the resource group.

-p
: Specifies a property of the resource group.

Pathprefix=/global/etc
: Specifies a directory into which the resources in the group can write administrative files.

Auto_start_on_new_cluster=False
: Specifies that the application resource group is not started automatically.

RG_dependencies=devgrp-stor-rg
: Specifies the resource groups that the application resource group depends on. In this example, the application resource group depends on the replication resource group.

  If the application resource group is switched over to a new primary node, the replication resource group is automatically switched over. However, if the replication resource group is switched over to a new primary node, the application resource group must be manually switched over.

nfs-rg
: The name of the application resource group.

5 **Add a** SUNW.HAStoragePlus **resource to the application resource group.**

```
nodeC# clresource create -g nfs-rg \
-t SUNW.HAStoragePlus \
-p FileSystemMountPoints=/global/mountpoint \
-p AffinityOn=True \
nfs-dg-rs
```

create
: Creates the resource.

-g
: Specifies the resource group to which the resource is added.

-t SUNW.HAStoragePlus
: Specifies that the resource is of the type SUNW.HAStoragePlus.

-p
: Specifies a property of the resource.

FileSystemMountPoints=/global/
: Specifies that the mount point for the file system is global.

AffinityOn=True
: Specifies that the application resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the application resource group fails over or is switched over, the associated device group is switched over.

```
nfs-dg-rs
```
The name of the HAStoragePlus resource for the NFS application.

For more information about these extension properties, see the SUNW.HAStoragePlus(5) man page.

6   Add a logical host name resource to the application resource group.

```
nodeC# clreslogicalhostname create -g nfs-rg \
lhost-nfsrg-sec
```

The logical host name of the application resource group on the secondary cluster is named lhost-nfsrg-sec.

7   Add an NFS resource to the application resource group.

```
nodeC# clresource create -g nfs-rg \
-t SUNW.nfs -p Resource_dependencies=nfs-dg-rs nfs-rg
```

8   Ensure that the application resource group does not come online on nodeC.

```
nodeC# clresource disable -n nodeC nfs-rs
nodeC# clresource disable -n nodeC nfs-dg-rs
nodeC# clresource disable -n nodeC lhost-nfsrg-sec
nodeC# clresourcegroup online -n "" nfs-rg
```

The resource group remains offline after a reboot, because *Auto_start_on_new_cluster=False*.

9   If the global volume is mounted on the primary cluster, unmount the global volume from the secondary cluster.

```
nodeC# umount /global/mountpoint
```

If the volume is mounted on a secondary cluster, the synchronization fails.

Next Steps   Go to "Example of How to Enable Data Replication" on page 103.

Example of How to Enable Data Replication

This section describes how data replication is enabled for the example configuration. This section uses the Sun StorageTek Availability Suite software commands sndradm and iiadm. For more information about these commands, see the Sun StorageTek Availability documentation.

This section contains the following procedures:

- "How to Enable Replication on the Primary Cluster" on page 104
- "How to Enable Replication on the Secondary Cluster" on page 106

▼ How to Enable Replication on the Primary Cluster

1 **Access** nodeA **as superuser or assume a role that provides** solaris.cluster.read **RBAC authorization.**

2 **Flush all transactions.**

```
nodeA# lockfs -a -f
```

3 **Confirm that the logical host names** lhost-reprg-prim **and** lhost-reprg-sec **are online.**

```
nodeA# clresourcegroup status
nodeC# clresourcegroup status
```

Examine the state field of the resource group.

4 **Enable remote mirror replication from the primary cluster to the secondary cluster.**

This step enables replication from the master volume on the primary cluster to the master volume on the secondary cluster. In addition, this step enables replication to the remote mirror bitmap on vol04.

- If the primary cluster and secondary cluster are unsynchronized, run this command:

    - For Sun StorEdge Availability Suite software:

    ```
    nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -e lhost-reprg-prim \
    /dev/vx/rdsk/devgrp/vol01 \
    /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
    /dev/vx/rdsk/devgrp/vol01 \
    /dev/vx/rdsk/devgrp/vol04 ip sync
    ```

    - For Sun StorageTek Availability Suite software:

    ```
    nodeA# /usr/sbin/sndradm -n -e lhost-reprg-prim \
    /dev/vx/rdsk/devgrp/vol01 \
    /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
    /dev/vx/rdsk/devgrp/vol01 \
    /dev/vx/rdsk/devgrp/vol04 ip sync
    ```

- If the primary cluster and secondary cluster are synchronized, run this command:

    - For Sun StorEdge Availability Suite software:

    ```
    nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -E lhost-reprg-prim \
    /dev/vx/rdsk/devgrp/vol01 \
    /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
    /dev/vx/rdsk/devgrp/vol01 \
    /dev/vx/rdsk/devgrp/vol04 ip sync
    ```

    - For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -E lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

5   **Enable autosynchronization.**

   ■ For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -a on lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

   ■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -a on lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

This step enables autosynchronization. When the active state of autosynchronization is set to on, the volume sets are resynchronized if the system reboots or a failure occurs.

6   **Verify that the cluster is in logging mode.**

   ■ For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -P
```

   ■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdsk/devgrp/vol01 ->
lhost-reprg-sec:/dev/vx/rdsk/devgrp/vol01
autosync: off, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devgrp, state: logging
```

In logging mode, the state is logging, and the active state of autosynchronization is off. When the data volume on the disk is written to, the bitmap file on the same disk is updated.

7   **Enable point-in-time snapshot.**

   ■ For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/iiadm -e ind \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
nodeA# /usr/opt/SUNWesm/sbin/iiadm -w \
/dev/vx/rdsk/devgrp/vol02
```

■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/iiadm -e ind \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
nodeA# /usr/sbin/iiadm -w \
/dev/vx/rdsk/devgrp/vol02
```

This step enables the master volume on the primary cluster to be copied to the shadow volume on the same cluster. The master volume, shadow volume, and point-in-time bitmap volume must be in the same device group. In this example, the master volume is vol01, the shadow volume is vol02, and the point-in-time bitmap volume is vol03.

8   Attach the point-in-time snapshot to the remote mirror set.

■ For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -I a \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
```

■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -I a \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
```

This step associates the point-in-time snapshot with the remote mirror volume set. Sun StorageTek Availability Suite software ensures that a point-in-time snapshot is taken before remote mirror replication can occur.

Next Steps   Go to

▼ How to Enable Replication on the Secondary Cluster

Before You Begin   Complete the procedure

1   Access nodeC **as superuser.**

2   Flush all transactions.

```
nodeC# lockfs -a -f
```

3   Enable remote mirror replication from the primary cluster to the secondary cluster.

- For Sun StorEdge Availability Suite software:

```
nodeC# /usr/opt/SUNWesm/sbin/sndradm -n -e lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

- For Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/sndradm -n -e lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

The primary cluster detects the presence of the secondary cluster and starts synchronization. Refer to the system log file /var/opt/SUNWesm/ds.log for Sun StorEdge Availability Suite or /var/adm for Sun StorageTek Availability Suite for information about the status of the clusters.

4   Enable independent point-in-time snapshot.

- For Sun StorEdge Availability Suite software:

```
nodeC# /usr/opt/SUNWesm/sbin/iiadm -e ind \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
nodeC# /usr/opt/SUNWesm/sbin/iiadm -w \
/dev/vx/rdsk/devgrp/vol02
```

- For Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/iiadm -e ind \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
nodeC# /usr/sbin/iiadm -w \
/dev/vx/rdsk/devgrp/vol02
```

5   Attach the point-in-time snapshot to the remote mirror set.

- For Sun StorEdge Availability Suite software:

```
nodeC# /usr/opt/SUNWesm/sbin/sndradm -I a \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
```

■ For Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/sndradm -I a \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol02 \
/dev/vx/rdsk/devgrp/vol03
```

Next Steps   Go to "Example of How to Perform Data Replication" on page 108.

Example of How to Perform Data Replication

This section describes how data replication is performed for the example configuration. This section uses the Sun StorageTek Availability Suite software commands sndradm and iiadm. For more information about these commands, see the Sun StorageTek Availability Suite documentation.

This section contains the following procedures:

■ "How to Perform a Remote Mirror Replication" on page 108
■ "How to Perform a Point-in-Time Snapshot" on page 110
■ "How to Verify That Replication Is Configured Correctly" on page 112

▼ How to Perform a Remote Mirror Replication

In this procedure, the master volume of the primary disk is replicated to the master volume on the secondary disk. The master volume is vol01 and the remote mirror bitmap volume is vol04.

1   Access nodeA **as superuser.**

2   Verify that the cluster is in logging mode.

■ For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -P
```

■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdsk/devgrp/vol01 ->
lhost-reprg-sec:/dev/vx/rdsk/devgrp/vol01
autosync: off, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devgrp, state: logging
```

In logging mode, the state is logging, and the active state of autosynchronization is off. When the data volume on the disk is written to, the bitmap file on the same disk is updated.

3   Flush all transactions.

```
nodeA# lockfs -a -f
```

4   Repeat Step 1 through Step 3 on nodeC.

5   Copy the master volume of nodeA to the master volume of nodeC.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -m lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

- For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -m lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

6   Wait until the replication is complete and the volumes are synchronized.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -w lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

- For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -w lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

7   Confirm that the cluster is in replicating mode.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -P
```

- For Sun StorageTek Availability Suite software:

  ```
  nodeA# /usr/sbin/sndradm -P
  ```

The output should resemble the following:

```
/dev/vx/rdsk/devgrp/vol01 ->
lhost-reprg-sec:/dev/vx/rdsk/devgrp/vol01
autosync: on, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devgrp, state: replicating
```

In replicating mode, the state is `replicating`, and the active state of autosynchronization is `on`. When the primary volume is written to, the secondary volume is updated by Sun StorageTek Availability Suite software.

Next Steps   Go to "How to Perform a Point-in-Time Snapshot" on page 110.

▼ How to Perform a Point-in-Time Snapshot

In this procedure, point-in-time snapshot is used to synchronize the shadow volume of the primary cluster to the master volume of the primary cluster. The master volume is `vol01`, the bitmap volume is `vol04`, and the shadow volume is `vol02`.

Before You Begin   Complete the procedure "How to Perform a Remote Mirror Replication" on page 108.

1   **Access** nodeA **as superuser or assume a role that provides** solaris.cluster.modify **and** solaris.cluster.admin **RBAC authorization.**

2   **Disable the resource that is running on** nodeA.

```
nodeA# clresource disable -n nodeA nfs-rs
```

3   **Change the primary cluster to logging mode.**

- For Sun StorEdge Availability Suite software:

  ```
  nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 ip sync
  ```

- For Sun StorageTek Availability Suite software:

  ```
  nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 ip sync
  ```

When the data volume on the disk is written to, the bitmap file on the same disk is updated. No replication occurs.

4 Synchronize the shadow volume of the primary cluster to the master volume of the primary cluster.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/iiadm -u s /dev/vx/rdsk/devgrp/vol02
nodeA# /usr/opt/SUNWesm/sbin/iiadm -w /dev/vx/rdsk/devgrp/vol02
```

- For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/iiadm -u s /dev/vx/rdsk/devgrp/vol02
nodeA# /usr/sbin/iiadm -w /dev/vx/rdsk/devgrp/vol02
```

5 Synchronize the shadow volume of the secondary cluster to the master volume of the secondary cluster.

- For Sun StorEdge Availability Suite software:

```
nodeC# /usr/opt/SUNWesm/sbin/iiadm -u s /dev/vx/rdsk/devgrp/vol02
nodeC# /usr/opt/SUNWesm/sbin/iiadm -w /dev/vx/rdsk/devgrp/vol02
```

- For Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/iiadm -u s /dev/vx/rdsk/devgrp/vol02
nodeC# /usr/sbin/iiadm -w /dev/vx/rdsk/devgrp/vol02
```

6 Restart the application on nodeA**.**

```
nodeA# clresource enable -n nodeA nfs-rs
```

7 Resynchronize the secondary volume with the primary volume.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -u lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

- For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

Next Steps Go to "How to Verify That Replication Is Configured Correctly" on page 112.

▼ How to Verify That Replication Is Configured Correctly

Before You Begin Complete the procedure "How to Perform a Point-in-Time Snapshot" on page 110.

1 Access nodeA **and** nodeC **as superuser or assume a role that provides** solaris.cluster.admin **RBAC authorization.**

2 **Verify that the primary cluster is in replicating mode, with autosynchronization on.**

- For Sun StorEdge Availability Suite software:

  nodeA# **/usr/opt/SUNWesm/sbin/sndradm -P**

- For Sun StorageTek Availability Suite software:

  nodeA# **/usr/sbin/sndradm -P**

The output should resemble the following:

```
/dev/vx/rdsk/devgrp/vol01 ->
lhost-reprg-sec:/dev/vx/rdsk/devgrp/vol01
autosync: on, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devgrp, state: replicating
```

In replicating mode, the state is replicating, and the active state of autosynchronization is on. When the primary volume is written to, the secondary volume is updated by Sun StorageTek Availability Suite software.

3 **If the primary cluster is not in replicating mode, put it into replicating mode.**

- For Sun StorEdge Availability Suite software:

  ```
  nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -u lhost-reprg-prim \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 ip sync
  ```

- For Sun StorageTek Availability Suite software:

  ```
  nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
  /dev/vx/rdsk/devgrp/vol01 \
  /dev/vx/rdsk/devgrp/vol04 ip sync
  ```

4   Create a directory on a client machine.

a.   Log in to a client machine as superuser.

You see a prompt that resembles the following:

client-machine#

b.   Create a directory on the client machine.

client-machine# **mkdir /***dir*

5   Mount the directory to the application on the primary cluster, and display the mounted directory.

a.   Mount the directory to the application on the primary cluster.

client-machine# **mount -o rw lhost-nfsrg-prim:/global/***mountpoint* **/***dir*

b.   Display the mounted directory.

client-machine# **ls /***dir*

6   Mount the directory to the application on the secondary cluster, and display the mounted directory.

a.   Unmount the directory from the application on the primary cluster.

client-machine# **umount /***dir*

b.   Take the application resource group offline on the primary cluster.

```
nodeA# clresource disable -n nodeA nfs-rs
nodeA# clresource disable -n nodeA nfs-dg-rs
nodeA# clresource disable -n nodeA lhost-nfsrg-prim
nodeA# clresourcegroup online -n "" nfs-rg
```

c.   Change the primary cluster to logging mode.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

- For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

When the data volume on the disk is written to, the bitmap file on the same disk is updated. No replication occurs.

d. Ensure that the `PathPrefix` **directory is available.**

```
nodeC# mount | grep /global/etc
```

e. Bring the application resource group online on the secondary cluster.

```
nodeC# clresourcegroup online -n nodeC nfs-rg
```

f. Access the client machine as superuser.

You see a prompt that resembles the following:

client-machine#

g. Mount the directory that was created in Step 4 to the application on the secondary cluster.

client-machine# `mount -o rw lhost-nfsrg-sec:/global/`*mountpoint* `/`*dir*

h. Display the mounted directory.

client-machine# `ls /`*dir*

7 Ensure that the directory displayed in Step 5 is the same as the directory displayed in Step 6.

8 Return the application on the primary cluster to the mounted directory.

a. Take the application resource group offline on the secondary cluster.

```
nodeC# clresource disable -n nodeC nfs-rs
nodeC# clresource disable -n nodeC nfs-dg-rs
nodeC# clresource disable -n nodeC lhost-nfsrg-sec
nodeC# clresourcegroup online -n "" nfs-rg
```

b. Ensure that the global volume is unmounted from the secondary cluster.

```
nodeC# umount /global/mountpoint
```

c. Bring the application resource group online on the primary cluster.

```
nodeA# clresourcegroup online -n nodeA nfs-rg
```

d. Change the primary cluster to replicating mode.

- For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -u lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

When the primary volume is written to, the secondary volume is updated by Sun StorageTek Availability Suite software.

See Also   "Example of How to Manage a Failover or Switchover" on page 115

Example of How to Manage a Failover or Switchover

This section describes how to provoke a switchover and how the application is transferred to the secondary cluster. After a switchover or failover, update the DNS entries. For additional information, see "Guidelines for Managing a Failover or Switchover" on page 86.

This section contains the following procedures:

■ "How to Provoke a Switchover" on page 115
■ "How to Update the DNS Entry" on page 117

▼ How to Provoke a Switchover

1 **Access** nodeA **and** nodeC **as superuser or assume a role that provides** solaris.cluster.admin **RBAC authorization.**

2 **Change the primary cluster to logging mode.**

   ■ For Sun StorEdge Availability Suite software:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -n -l lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

   ■ For Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devgrp/vol01 \
/dev/vx/rdsk/devgrp/vol04 ip sync
```

When the data volume on the disk is written to, the bitmap volume on the same device group is updated. No replication occurs.

3   **Confirm that the primary cluster and the secondary cluster are in logging mode, with autosynchronization off.**

   a. **On** `nodeA`**, confirm the mode and setting:**

      ■ For Sun StorEdge Availability Suite software:

         nodeA# **`/usr/opt/SUNWesm/sbin/sndradm -P`**

      ■ For Sun StorageTek Availability Suite software:

         nodeA# **`/usr/sbin/sndradm -P`**

      The output should resemble the following:

      ```
      /dev/vx/rdsk/devgrp/vol01 ->
      lhost-reprg-sec:/dev/vx/rdsk/devgrp/vol01
      autosync:off, max q writes:4194304,max q fbas:16384,mode:sync,ctag:
      devgrp, state: logging
      ```

   b. **On** `nodeC`**, confirm the mode and setting:**

      ■ For Sun StorEdge Availability Suite software:

         nodeC# **`/usr/opt/SUNWesm/sbin/sndradm -P`**

      ■ For Sun StorageTek Availability Suite software:

         nodeC# **`/usr/sbin/sndradm -P`**

      The output should resemble the following:

      ```
      /dev/vx/rdsk/devgrp/vol01 <-
      lhost-reprg-prim:/dev/vx/rdsk/devgrp/vol01
      autosync:off, max q writes:4194304,max q fbas:16384,mode:sync,ctag:
      devgrp, state: logging
      ```

   For nodeA and nodeC, the state should be logging, and the active state of autosynchronization should be off.

4   **Confirm that the secondary cluster is ready to take over from the primary cluster.**

   nodeC# **`fsck -y /dev/vx/rdsk/devgrp/vol01`**

5   **Switch over to the secondary cluster.**

   nodeC# **`clresourcegroup switch -n nodeC nfs-rg`**

Next Steps Go to "How to Update the DNS Entry" on page 117.

▼ How to Update the DNS Entry

For an illustration of how DNS maps a client to a cluster, see Figure 4–8.

Before You Begin Complete the procedure "How to Provoke a Switchover" on page 115.

1 Start the nsupdate **command.**

For information, see the nsupdate(1M) man page.

2 Remove the current DNS mapping between the logical host name of the application resource group and the cluster IP address, for both clusters.

> **update delete lhost-nfsrg-prim A**
> **update delete lhost-nfsrg-sec A**
> **update delete** *ipaddress1rev***.in-addr.arpa** *ttl* **PTR lhost-nfsrg-prim**
> **update delete** *ipaddress2rev***.in-addr.arpa** *ttl* **PTR lhost-nfsrg-sec**

ipaddress1rev	The IP address of the primary cluster, in reverse order.
ipaddress2rev	The IP address of the secondary cluster, in reverse order.
ttl	The time to live, in seconds. A typical value is 3600.

3 Create a new DNS mapping between the logical host name of the application resource group and the cluster IP address, for both clusters.

Map the primary logical host name to the IP address of the secondary cluster and map the secondary logical host name to the IP address of the primary cluster.

> **update add lhost-nfsrg-prim** *ttl* **A** *ipaddress2fwd*
> **update add lhost-nfsrg-sec** *ttl* **A** *ipaddress1fwd*
> **update add** *ipaddress2rev***.in-addr.arpa** *ttl* **PTR lhost-nfsrg-prim**
> **update add** *ipaddress1rev***.in-addr.arpa** *ttl* **PTR lhost-nfsrg-sec**

ipaddress2fwd	The IP address of the secondary cluster, in forward order.
ipaddress1fwd	The IP address of the primary cluster, in forward order.

5

Administering Global Devices, Disk-Path Monitoring, and Cluster File Systems

This chapter provides information about and procedures for administering global devices, disk-path monitoring and cluster file systems.

- "Overview of Administering Global Devices and the Global Namespace" on page 119
- "Administering Storage-Based Replicated Devices" on page 122
- "Overview of the Administering Cluster File Systems" on page 133
- "Administering Device Groups" on page 135
- "Administering the SCSI Protocol Settings for Storage Devices" on page 169
- "Administering Cluster File Systems" on page 173
- "Administering Disk-Path Monitoring" on page 179

For a high-level description of the related procedures in this chapter, see Table 5–3.

For conceptual information related to global devices, the global namespace, device groups, disk-path monitoring, and the cluster file system, see *Sun Cluster Concepts Guide for Solaris OS*.

Overview of Administering Global Devices and the Global Namespace

Administration of Sun Cluster device groups depends on the volume manager that is installed on the cluster. Solaris Volume Manager is "cluster-aware," so you add, register, and remove device groups by using the Solaris Volume Manager metaset(1M) command. If you are using VERITAS Volume Manager (VxVM), you create disk groups by using VxVM commands. You register the disk groups as Sun Cluster device groups with the clsetup utility. When removing VxVM device groups, you use both the clsetup command and VxVM commands.

Note – For the Solaris 10 OS, global devices are not directly accessible from non-global zones.

Sun Cluster software automatically creates a raw-disk device group for each disk and tape device in the cluster. However, cluster device groups remain in an offline state until you access the groups as

global devices. When administering device groups, or volume manager disk groups, you need to be on the cluster node that is the primary node for the group.

Normally, you do not need to administer the global device namespace. The global namespace is automatically set up during installation and automatically updated during Solaris OS reboots. However, if the global namespace needs to be updated, you can run the `cldevice populate` command from any cluster node. This command causes the global namespace to be updated on all other cluster node members, as well as on nodes that might join the cluster in the future.

Global Device Permissions for Solaris Volume Manager

Changes made to global device permissions are not automatically propagated to all the nodes in the cluster for Solaris Volume Manager and disk devices. If you want to change permissions on global devices, you must manually change the permissions on all the nodes in the cluster. For example, if you want to change permissions on global device `/dev/global/dsk/d3s0` to 644, you must issue the following command on all nodes in the cluster:

```
# chmod 644 /dev/global/dsk/d3s0
```

VxVM does not support the `chmod` command. To change global device permissions in VxVM, consult the VxVM administrator's guide.

Dynamic Reconfiguration With Global Devices

You must consider the following issues when completing dynamic reconfiguration (DR) operations on disk and tape devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support. The only exception is for the operating system quiescence operation. Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect nonnetwork IO devices during a DR detach operation.

- Sun Cluster rejects DR remove-board operations on active devices in the primary node. DR operations can be performed on inactive devices in the primary node and on any devices in the secondary nodes.

- After the DR operation, cluster data access continues as before.

- Sun Cluster rejects DR operations that impact the availability of quorum devices. See "Dynamic Reconfiguration With Quorum Devices" on page 189 for more information.

> ⚠️ **Caution** – If the current primary node fails while you are performing the DR operation on a secondary node, cluster availability is impacted. The primary node will have no place to fail over until a new secondary node is provided.

To perform DR operations on global devices, complete the following steps in the order indicated.

TABLE 5–1 Task Map: Dynamic Reconfiguration With Disk and Tape Devices

Task	For Instructions
1. If a DR operation that affects an active device group must be performed on the current primary node, switch the primary and secondary nodes before performing the DR remove operation on the device.	"How to Switch the Primary for a Device Group" on page 166
2. Perform the DR removal operation on the device being removed.	*Sun Enterprise 10000 DR Configuration Guide* and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* in the *Solaris 9 on Sun Hardware*, and *Solaris 10 on Sun Hardware* collections.

SPARC: VERITAS Volume Manager Administration Considerations

- For Sun Cluster to maintain the VxVM namespace, you must register any VxVM disk group or volume changes as Sun Cluster device group configuration changes. Registering these changes ensures that the namespace on all cluster nodes is updated. Examples of configuration changes that impact the namespace include adding, removing, or renaming a volume. Changing the volume permissions, owner, or group ID also impacts the namespace.

  > **Note** – Never import or deport VxVM disk groups by using VxVM commands after the disk group has been registered with the cluster as a Sun Cluster device group. The Sun Cluster software handles all cases where disk groups need to be imported or be deported.

- Each VxVM disk group must have a cluster-wide unique minor number. By default, when a disk group is created, VxVM chooses a random number that is a multiple of 1000 as that disk group's base minor number. For most configurations with only a small number of disk groups, the minor number is sufficient to guarantee uniqueness. The minor number for a newly created disk group might conflict with the minor number of a pre-existing disk group that was imported on a different node. In this case, attempting to register the Sun Cluster device group fails. To fix this problem, the new disk group should be given a new minor number that is a unique value and then registered as a Sun Cluster device group.

- If you are setting up a mirrored volume, Dirty Region Logging (DRL) can be used to decrease volume recovery time after a node failure. Use of DRL is strongly recommended, although use of DRL could decrease I/O throughput.

- VxVM does not support the chmod command. To change global device permissions in VxVM, consult the VxVM administrator's guide.

- Sun Cluster 3.2 software does not support the VxVM Dynamic Multipathing (DMP) management of multiple paths from the same node.

- If you use VxVM to set up shared disk groups for Oracle Parallel Server or Oracle RAC, use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*. Creating shared disk groups for Oracle Parallel Server or Oracle RAC differs from creating other disk groups. You must import the Oracle Parallel Server or Oracle RAC shared disk groups by using vxdg -s. You do not register the Oracle Parallel Server or Oracle RAC shared disk groups with the cluster framework. To create other VxVM disk groups, see "SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 144.

Administering Storage-Based Replicated Devices

You can configure a Sun Cluster device group to contain devices that are replicated by using storage-based replication. Sun Cluster supports Hitachi TrueCopy software for storage-based replication.

Before you can replicate data with Hitachi TrueCopy software, you must be familiar with the Hitachi TrueCopy documentation and have the Hitachi TrueCopy product and the latest Hitachi TrueCopy patches installed on your system. For information about installing the Hitachi TrueCopy software, see the Hitachi TrueCopy product documentation.

The Hitachi TrueCopy software configures a pair of devices as replicas with one device as the primary replica and the other device as the secondary replica. At any given time, the device attached to one set of nodes will be the primary replicas. The device attached to the other set of nodes will be the secondary replica.

In Sun Cluster, the primary replica is automatically moved whenever the Sun Cluster device group to which the replica belongs is moved. Therefore, the replica primary should never be moved in Sun Cluster by issuing a horcmtakeover directly. Rather, the takeover should be accomplished by moving the associated Sun Cluster device group.

The following table lists the tasks you must perform to set up a storage-based replicated device.

TABLE 5–2 Task Map: Administering a Storage-Based Replicate Device

Task	Instructions
Install the TrueCopy software on your storage device and nodes.	The documentation that shipped with your Hitachi storage device.
Configure the Hitachi replication group.	"How to Configure a Hitachi TrueCopy Replication Group" on page 123

TABLE 5–2 Task Map: Administering a Storage-Based Replicate Device   *(Continued)*

Task	Instructions
Configure the DID device.	"How to Configure DID Devices for Replication" on page 124
Register the replicated group.	"How to Add and Register a Device Group (Solaris Volume Manager)" on page 138
Verify the configuration.	"How to Verify Replicated Global Device Group Configuration" on page 126

▼ How to Configure a Hitachi TrueCopy Replication Group

Before You Begin First, configure the Hitachi TrueCopy device groups on shared disks in the primary cluster. This configuration information is specified in the/etc/horcm.conf file on each of the cluster's nodes that has access to the Hitachi array. For more information about how to configure the /etc/horcm.conf file, see the *Sun StorEdge SE 9900 V Series Command and Control Interface User and Reference Guide.*

1 **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on the node connected to the storage array.**

2 **Add the** horcm **entry to the** /etc/services **file.**

```
horcm  9970/udp
```

Specify a port number and protocol name for the new entry.

3 **Specify the Hitachi TrueCopy device group configuration information in the** /etc/horcm.conf **file.**

For instructions, refer to the documentation that shipped with your TrueCopy software.

4 **Start the TrueCopy CCI daemon by running the** horcmstart.sh **command on all nodes.**

```
# /usr/bin/horcmstart.sh
```

5 **If you have not already created the replica pairs, create them now.**

Use the paircreate command to create your replica pairs with the desired fence level. For instructions on creating the replica pairs, refer to your TrueCopy documentation.

6 **On each node configured with replicated devices, verify that data replication is set up correctly by using the** pairdisplay **command.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321   58..P-VOL PAIR DATA ,12345 29   -
group-name pair1(R) (CL1-A , 0, 29)12345   29..S-VOL PAIR DATA ,----- 58   -
```

7 Verify that all nodes can master the replication groups.

a. Determine which node contains the primary replica and which node contains the secondary replica by using the `pairdisplay` **command.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321  58..P-VOL PAIR DATA ,12345 29   -
group-name pair1(R) (CL1-A , 0, 29)12345   29..S-VOL PAIR DATA ,----- 58   -
```

The node with the local (L) device in the P-VOL state contains the primary replica and the node with the local (L) device in the S-VOL state contains the secondary replica.

b. Make the secondary node the master by running the `horctakeover` **command on the node that contains the secondary replica.**

```
# horctakeover -g group-name
```

Wait for the initial data copy to complete before proceeding to the next step.

c. Verify that the node that performed the `horctakeover` **now has the local (L) device in the** P-VOL **state.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321  58..S-VOL PAIR DATA ,12345 29   -
group-name pair1(R) (CL1-A , 0, 29)12345   29..P-VOL PAIR DATA ,----- 58   -
```

d. Run the `horctakeover` **command on the node that originally contained the primary replica.**

```
# horctakeover -g group-name
```

e. Verify that the primary node has changed back to the original configuration by running the `pairdisplay` **command.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321  58..P-VOL PAIR DATA ,12345 29   -
group-name pair1(R) (CL1-A , 0, 29)12345   29..S-VOL PAIR DATA ,----- 58   -
```

Next Steps   Continue the configuration of your replicated device by following the instructions in "How to Configure DID Devices for Replication" on page 124.

▼ How to Configure DID Devices for Replication

Before You Begin   After you have configured a device group for your replicated device, you must configure the device identifier (DID) driver that the replicated device uses.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1. **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on any node of the cluster.**

2. **Confirm that the** `horcm` **daemon is running on all nodes.**

   `# /usr/bin/horcmstart.sh`

3. **Determine which node contains the secondary replica by running the** `pairdisplay` **command.**

   ```
   # pairdisplay -g group-name
   Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
   group-name pair1(L) (CL1-C , 0, 9) 54321  58..P-VOL PAIR DATA ,12345 29   -
   group-name pair1(R) (CL1-A , 0, 29)12345  29..S-VOL PAIR DATA ,----- 58   -
   ```

   The node with the local (L) device in the `S-VOL` state contains the secondary replica.

4. **On the node with secondary replica (as determined by the previous step), configure the DID devices for use with storage-based replication.**

   This command combines the two separate DID instances for the device replica pairs into a single, logical DID instance. The single instance enables the device to be used by volume management software from both nodes.

   ⚠ **Caution –** If multiple nodes are connected to the secondary replica, run this command on only one of these nodes.

   `# cldevice replicate -D` *primary-replica-nodename* `-s secondary replica`

   primary-replica-nodename
     Specifies the name of the remote node that contains the primary replica.

   `-S`
     Specifies a source node other than the current node.

   Note – By default, the current node is the source node. Use the `-S` option to specify a different source node.

5. **Verify that the DID instances have been combined.**

   `# cldevice list -v` *device*

6. **Verify that the TrueCopy replication is set.**

   `# cldevice show` *device* |

7. **If the DID remapping did not successfully combine all replicated devices, combine the individual replicated devices manually.**

⚠ **Caution** – Exercise extreme care when combining DID instances manually. Improper device remapping can cause data corruption.

a. On the node that contains the secondary replica, run the `cldevice combine` **command. This node must be local.**

`# cldevice combine -d` *destination-instance* *source-instance*

`-d` *destination-instance*	The remote DID instance, which corresponds to the primary replica.
source-instance	The local DID instance, which corresponds to the secondary replica.

b. Verify that the DID remapping occurred successfully.

`# cldevice list` *desination-instance* *source-instance*

One of the DID instances should not be listed.

8 On all nodes, verify that the DID devices for all combined DID instances are accessible.

`# cldevice list -v`

Next Steps To complete the configuration of your replicated device group, perform the steps in the following procedures.

- "How to Add and Register a Device Group (Solaris Volume Manager)" on page 138 or "SPARC: How to Register a Disk Group as a Device Group (VERITAS Volume Manager)" on page 149

  When registering the device group, make sure to give it the same name as the TrueCopy replication group.

- "How to Verify Replicated Global Device Group Configuration" on page 126

▼ How to Verify Replicated Global Device Group Configuration

Before You Begin Before you verify the global device group, you must first create them. For information about creating a Solstice DiskSuite or Solaris Volume Manager device group, see "How to Add and Register a Device Group (Solaris Volume Manager)" on page 138. For information about creating a VERITAS Volume Manager device group, see "SPARC: How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 145.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Verify that the primary device group corresponds to the same node as the node that contains the primary replica.**

    # **pairdisplay -g** *group-name*
    # **cldevicegroup status -n** *nodename group-name*

2   **Perform a trial switchover to ensure that the device groups are configured correctly and the replicas can move between nodes.**

    If the device group is offline, bring it online.

    # **cldevicegroup switch -n** *nodename group-name*

    -n *nodename*   The node to which the device group is switched. This node becomes the new primary

3   **Verify that the switchover was successful by comparing the output of the following commands.**

    # **pairdisplay -g** *group-name*
    # **cldevicegroup status -n** *nodename group-name*

Example: Configuring a TrueCopy Replication Group for Sun Cluster

This example completes the Sun Cluster specific steps necessary to set up TrueCopy replication in your cluster. The example assumes that you have already performed the following tasks:

- Set up your Hitachi LUNs
- Installed the TrueCopy software on your storage device and cluster nodes
- Configured the replication pairs on your cluster nodes

  For instructions about configuring your replication pairs, see "How to Configure a Hitachi TrueCopy Replication Group" on page 123.

This example involves a three-node cluster that uses TrueCopy. The cluster is spread across two remote sites, with two nodes at one site and one node at the other site. Each site has its own Hitachi storage device.

The following examples show the TrueCopy /etc/horcm.conf configuration file on each node.

EXAMPLE 5–1 TrueCopy Configuration File on Node 1

```
HORCM_DEV
#dev_group      dev_name      port#         TargetID      LU#         MU#
VG01            pair1         CL1-A         0             29
VG01            pair2         CL1-A         0             30
VG01            pair3         CL1-A         0             31
HORCM_INST
```

EXAMPLE 5–1 TrueCopy Configuration File on Node 1   *(Continued)*

```
#dev_group      ip_address    service
VG01            node-3        horcm
```

EXAMPLE 5–2 TrueCopy Configuration File on Node 2

```
HORCM_DEV
#dev_group      dev_name      port#       TargetID      LU#        MU#
VG01            pair1         CL1-A       0             29
VG01            pair2         CL1-A       0             30
VG01            pair3         CL1-A       0             31
HORCM_INST
#dev_group      ip_address    service
VG01            node-3        horcm
```

EXAMPLE 5–3 TrueCopy Configuration File on Node 3

```
HORCM_DEV
#dev_group      dev_name      port#       TargetID      LU#        MU#
VG01            pair1         CL1-C       0             09
VG01            pair2         CL1-C       0             10
VG01            pair3         CL1-C       0             11
HORCM_INST
#dev_group      ip_address    service
VG01            node-1        horcm
VG01            node-2        horcm
```

In the preceding examples, three LUNs are replicated between the two sites. The LUNs are all in a replication group named VG01. The pairdisplay command verifies this information and shows that Node 3 has the primary replica.

EXAMPLE 5–4 pairdisplay Command Output on Node 1

```
# pairdisplay -g VG01
Group   PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
VG01    pair1(L)     (CL1-A , 0, 29)61114   29..S-VOL PAIR DATA ,-----   58 -
VG01    pair1(R)     (CL1-C , 0,  9)20064   58..P-VOL PAIR DATA ,61114   29 -
VG01    pair2(L)     (CL1-A , 0, 30)61114   30..S-VOL PAIR DATA ,-----   59 -
VG01    pair2(R)     (CL1-C , 0, 10)20064   59..P-VOL PAIR DATA ,61114   30 -
VG01    pair3(L)     (CL1-A , 0, 31)61114   31..S-VOL PAIR DATA ,-----   60 -
VG01    pair3(R)     (CL1-C , 0, 11)20064   60..P-VOL PAIR DATA ,61114   31 -
```

EXAMPLE 5–5 `pairdisplay` Command Output on Node 2

```
# pairdisplay -g VG01
Group   PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
VG01    pair1(L)     (CL1-A , 0, 29)61114    29..S-VOL PAIR DATA  ,-----    58  -
VG01    pair1(R)     (CL1-C , 0,  9)20064    58..P-VOL PAIR DATA  ,61114    29  -
VG01    pair2(L)     (CL1-A , 0, 30)61114    30..S-VOL PAIR DATA  ,-----    59  -
VG01    pair2(R)     (CL1-C , 0, 10)20064    59..P-VOL PAIR DATA  ,61114    30  -
VG01    pair3(L)     (CL1-A , 0, 31)61114    31..S-VOL PAIR DATA  ,-----    60  -
VG01    pair3(R)     (CL1-C , 0, 11)20064    60..P-VOL PAIR DATA  ,61114    31  -
```

EXAMPLE 5–6 `pairdisplay` Command Output on Node 3

```
# pairdisplay -g VG01
Group   PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
VG01    pair1(L)     (CL1-C , 0,  9)20064    58..P-VOL PAIR DATA  ,61114    29  -
VG01    pair1(R)     (CL1-A , 0, 29)61114    29..S-VOL PAIR DATA  ,-----    58  -
VG01    pair2(L)     (CL1-C , 0, 10)20064    59..P-VOL PAIR DATA  ,61114    30  -
VG01    pair2(R)     (CL1-A , 0, 30)61114    30..S-VOL PAIR DATA  ,-----    59  -
VG01    pair3(L)     (CL1-C , 0, 11)20064    60..P-VOL PAIR DATA  ,61114    31  -
VG01    pair3(R)     (CL1-A , 0, 31)61114    31..S-VOL PAIR DATA  ,-----    60  -
```

To see which disks are being used, use the `-fd` option of the `pairdisplay` command as shown in the following examples.

EXAMPLE 5–7 `pairdisplay` Command Output on Node 1, Showing Disks Used

```
# pairdisplay -fd -g VG01
Group PairVol(L/R) Device_File                       ,Seq#,LDEV#.P/S,Status,Fence,Seq#,P-LDEV# M
VG01 pair1(L) c6t500060E8000000000000EEBA0000001Dd0s2 61114 29..S-VOL PAIR DATA ,-----    58  -
VG01 pair1(R) c5t50060E800000000000004E600000003Ad0s2 20064 58..P-VOL PAIR DATA ,61114    29  -
VG01 pair2(L) c6t500060E8000000000000EEBA0000001Ed0s2 61114 30..S-VOL PAIR DATA ,-----    59  -
VG01 pair2(R) c5t50060E800000000000004E600000003Bd0s2 0064  59..P-VOL PAIR DATA ,61114    30  -
VG01 pair3(L) c6t500060E8000000000000EEBA0000001Fd0s2 61114 31..S-VOL PAIR DATA ,-----    60  -
VG01 pair3(R) c5t50060E800000000000004E600000003Cd0s2 20064 60..P-VOL PAIR DATA ,61114    31  -
```

EXAMPLE 5–8 `pairdisplay` Command Output on Node 2, Showing Disks Used

```
# pairdisplay -fd -g VG01
Group PairVol(L/R) Device_File                       ,Seq#,LDEV#.P/S,Status,Fence,Seq#,P-LDEV# M
VG01 pair1(L) c5t500060E8000000000000EEBA0000001Dd0s2 61114 29..S-VOL PAIR DATA ,-----    58  -
VG01 pair1(R) c5t50060E800000000000004E600000003Ad0s2 20064 58..P-VOL PAIR DATA ,61114    29  -
VG01 pair2(L) c5t500060E8000000000000EEBA0000001Ed0s2 61114 30..S-VOL PAIR DATA ,-----    59  -
VG01 pair2(R) c5t50060E800000000000004E600000003Bd0s2 20064 59..P-VOL PAIR DATA ,61114    30  -
VG01 pair3(L) c5t500060E8000000000000EEBA0000001Fd0s2 61114 31..S-VOL PAIR DATA ,-----    60  -
```

EXAMPLE 5–8 `pairdisplay` Command Output on Node 2, Showing Disks Used   *(Continued)*

```
VG01 pair3(R) c5t50060E800000000000004E600000003Cd0s2 20064 60..P-VOL PAIR DATA  ,61114   31  -
```

EXAMPLE 5–9 `pairdisplay` Command Output on Node 3, Showing Disks Used

```
# pairdisplay -fd -g VG01
Group PairVol(L/R) Device_File                     ,Seq#,LDEV#.P/S,Status,Fence ,Seq#,P-LDEV# M
VG01 pair1(L) c5t50060E800000000000004E600000003Ad0s2 20064  58..P-VOL PAIR DATA ,61114   29 -
VG01 pair1(R) c6t500060E8000000000000EEBA0000001Dd0s2 61114  29..S-VOL PAIR DATA ,-----   58 -
VG01 pair2(L) c5t50060E800000000000004E600000003Bd0s2 20064  59..P-VOL PAIR DATA ,61114   30 -
VG01 pair2(R) c6t500060E8000000000000EEBA0000001Ed0s2 61114  30..S-VOL PAIR DATA ,-----   59 -
VG01 pair3(L) c5t50060E800000000000004E600000003Cd0s2 20064  60..P-VOL PAIR DATA ,61114   31 -
VG01 pair3(R) c6t500060E8000000000000EEBA0000001Fd0s2 61114  31..S-VOL PAIR DATA ,-----   60 -
```

These examples show that the following disks are being used:

- On Node 1:
  - c6t500060E8000000000000EEBA0000001Dd0s2
  - c6t500060E8000000000000EEBA0000001Ed0s2
  - c6t500060E8000000000000EEBA0000001Fd0s
- On Node 2:
  - c5t500060E8000000000000EEBA0000001Dd0s2
  - c5t500060E8000000000000EEBA0000001Ed0s2
  - c5t500060E8000000000000EEBA0000001Fd0s2
- On Node 3:
  - c5t50060E800000000000004E600000003Ad0s2
  - c5t50060E800000000000004E600000003Bd0s2
  - c5t50060E800000000000004E600000003Cd0s2

To see the DID devices that corresponds to these disks, use the `cldevice list` command as shown in the following examples.

EXAMPLE 5–10 Displaying DIDs Corresponding to the Disks Used

```
# cldevice list -v

DID Device  Full Device Path
----------  ----------------
1           node-1:/dev/rdsk/c0t0d0  /dev/did/rdsk/d1
2           node-1:/dev/rdsk/c0t6d0  /dev/did/rdsk/d2
11          node-1:/dev/rdsk/c6t500060E8000000000000EEBA00000020d0 /dev/did/rdsk/d11
11          node-2:/dev/rdsk/c5t500060E8000000000000EEBA00000020d0 /dev/did/rdsk/d11
12          node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Fd0 /dev/did/rdsk/d12
```

EXAMPLE 5–10 Displaying DIDs Corresponding to the Disks Used   *(Continued)*

12	**node-2:/dev/rdsk/c5t500060E8000000000000EEBA0000001Fd0**	**/dev/did/rdsk/d12**
13	**node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Ed0**	**/dev/did/rdsk/d13**
13	**node-2:/dev/rdsk/c5t500060E8000000000000EEBA0000001Ed0**	**/dev/did/rdsk/d13**
14	**node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Dd0**	**/dev/did/rdsk/d14**
14	**node-2:/dev/rdsk/c5t500060E8000000000000EEBA0000001Dd0**	**/dev/did/rdsk/d14**
18	node-3:/dev/rdsk/c0t0d0	/dev/did/rdsk/d18
19	node-3:/dev/rdsk/c0t6d0	/dev/did/rdsk/d19
20	node-3:/dev/rdsk/c5t50060E800000000000004E6000000013d0	/dev/did/rdsk/d20
21	node-3:/dev/rdsk/c5t50060E800000000000004E600000003Dd0	/dev/did/rdsk/d21
22	**node-3:/dev/rdsk/c5t50060E800000000000004E600000003Cd0**	**/dev/did/rdsk/d2223**
23	**node-3:/dev/rdsk/c5t50060E800000000000004E600000003Bd0**	**/dev/did/rdsk/d23**
24	**node-3:/dev/rdsk/c5t50060E800000000000004E600000003Ad0**	**/dev/did/rdsk/d24**

When combining the DID instances for each pair of replicated devices, cldevice list should combine DID instance 12 with 22, instance 13 with 23 and instance 14 with 24. Because Node 3 has the primary replica, run the cldevice -T command from either Node 1 or Node 2. Always combine the instances from a node that has the secondary replica. Run this command from a single node only, not on both nodes.

The following example shows the output when combining DID instances by running the command on Node 1.

EXAMPLE 5–11 Combining DID Instances

```
# cldevice replicate -D node-3
Remapping instances for devices replicated with node-3...
VG01 pair1 L node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Dd0
VG01 pair1 R node-3:/dev/rdsk/c5t50060E800000000000004E600000003Ad0
Combining instance 14 with 24
VG01 pair2 L node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Ed0
VG01 pair2 R node-3:/dev/rdsk/c5t50060E800000000000004E600000003Bd0
Combining instance 13 with 23
VG01 pair3 L node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Fd0
VG01 pair3 R node-3:/dev/rdsk/c5t50060E800000000000004E600000003Cd0
Combining instance 12 with 22
```

Checking the cldevice list output, the LUNs from both sites now have the same DID instance. Having the same DID instance makes each replica pair look like a single DID device, as the following example shows.

EXAMPLE 5–12 Displaying the Combined DIDs

```
# cldevice list -v
DID Device  Full Device Path
```

EXAMPLE 5–12 Displaying the Combined DIDs   *(Continued)*

```
----------   ----------------
1            node-1:/dev/rdsk/c0t0d0  /dev/did/rdsk/d1
2            node-1:/dev/rdsk/c0t6d0  /dev/did/rdsk/d2
11           node-1:/dev/rdsk/c6t500060E8000000000000EEBA00000020d0 /dev/did/rdsk/d11
11           node-2:/dev/rdsk/c5t500060E8000000000000EEBA00000020d0 /dev/did/rdsk/d11
18           node-3:/dev/rdsk/c0t0d0  /dev/did/rdsk/d18
19           node-3:/dev/rdsk/c0t6d0  /dev/did/rdsk/d19
20           node-3:/dev/rdsk/c5t50060E800000000000004E6000000013d0 /dev/did/rdsk/d20
21           node-3:/dev/rdsk/c5t50060E800000000000004E600000003Dd0 /dev/did/rdsk/d21
```
22   node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Fd0 /dev/did/rdsk/d1222
22   node-2:/dev/rdsk/c5t500060E8000000000000EEBA0000001Fd0 /dev/did/rdsk/d12
22   node-3:/dev/rdsk/c5t50060E800000000000004E600000003Cd0 /dev/did/rdsk/d22
23   node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Ed0 /dev/did/rdsk/d13
23   node-2:/dev/rdsk/c5t500060E8000000000000EEBA0000001Ed0 /dev/did/rdsk/d13
23   node-3:/dev/rdsk/c5t50060E800000000000004E600000003Bd0 /dev/did/rdsk/d23
24   node-1:/dev/rdsk/c6t500060E8000000000000EEBA0000001Dd0 /dev/did/rdsk/d24
24   node-2:/dev/rdsk/c5t500060E8000000000000EEBA0000001Dd0 /dev/did/rdsk/d24
24   node-3:/dev/rdsk/c5t50060E800000000000004E600000003Ad0 /dev/did/rdsk/d24

The next step is to create the volume manager device group. Issue this command from the node that has the primary replica, in this example Node 3. Give the device group the same name as the replica group, as the following example shows.

EXAMPLE 5–13 Creating the Solaris Volume Manager Device Group

```
# metaset -s VG01 -ah phys-deneb-3
# metaset -s VG01 -ah phys-deneb-1
# metaset -s VG01 -ah phys-deneb-2
# metaset -s VG01 -a /dev/did/rdsk/d22
# metaset -s VG01 -a /dev/did/rdsk/d23
# metaset -s VG01 -a /dev/did/rdsk/d24
# metaset
Set name = VG01, Set number = 1

Host              Owner
  phys-deneb-3      Yes
  phys-deneb-1
  phys-deneb-2

Drive Dbase
d22   Yes
d23   Yes
d24   Yes
```

At this point the device group is usable, metadevices can be created, and the device group can be moved to any of the three nodes. However, to make switchovers and failovers more efficient, run cldevicegroup set to mark the device group as replicated in cluster configuration.

EXAMPLE 5–14 Making Switchovers and Failovers Efficient

```
# cldevicegroup sync VG01
# cldevicegroup show VG01
=== Device Groups===

Device Group Name                       VG01
  Type:                                 SVM
  failback:                             no
  Node List:                            phys-deneb-3, phys-deneb-1, phys-deneb-2
  preferenced:                          yes
  numsecondaries:                       1
  device names:                         VG01
  Replication type:                     truecopy
```

Configuration of the replication group is complete with this step. To verify that the configuration was successful, perform the steps in "How to Verify Replicated Global Device Group Configuration" on page 126.

Overview of the Administering Cluster File Systems

No special Sun Cluster commands are necessary for cluster file system administration. Administer a cluster file system as you would any other Solaris file system, using standard Solaris file-system commands, such as mount and newfs. Mount cluster file systems by specifying the -g option to the mount command. Cluster file systems can also be automatically mounted at boot. Cluster file systems are only visible from the global zone. If you require the cluster file system data to be accessible from a non-global zone, map the data to the non-global zone with zoneadm(1M)zoneadm or HAStoragePlus.

Note – When the cluster file system reads files, the file system does not update the access time on those files.

Cluster File-System Restrictions

The following restrictions apply to the cluster file-system administration:

- The unlink(1M) command is not supported on nonempty directories.
- The lockfs -d command is not supported. Use lockfs -n as a workaround.

- You cannot remount a cluster file system with the `directio` mount option added at remount time.

- You cannot set the `directio` mount option on a single file by using the `directio` ioctl.

SPARC: Guidelines to Support VxFS

The following VxFS features are not supported in a Sun Cluster 3.2 cluster file system. They are, however, supported in a local file system.

- Quick I/O
- Snapshots
- Storage checkpoints
- VxFS-specific mount options:
    - `convosync` (Convert O_SYNC)
    - `mincache`
    - `qlog`, `delaylog`, `tmplog`
- VERITAS cluster file system (requires VxVM cluster feature & VERITAS Cluster Server)

Cache advisories can be used, but the effect is observed on the given node only.

All other VxFS features and options that are supported in a cluster file system are supported by Sun Cluster 3.2 software. See VxFS documentation for details about VxFS options that are supported in a cluster configuration.

The following guidelines for using VxFS to create highly available cluster file systems are specific to a Sun Cluster 3.2 configuration.

- Create a VxFS file system by the following procedures in the VxFS documentation.

- Mount and unmount a VxFS file system from the primary node. The primary node masters the disk on which the VxFS file system resides. A VxFS file system mount or unmount operation that is performed from a secondary node might fail.

- Perform all VxFS administration commands from the primary node of the VxFS cluster file system.

The following guidelines for administering VxFS cluster file systems are not specific to Sun Cluster 3.2 software. However, the guidelines are different from the way you administer UFS cluster file systems.

- You can administer files on a VxFS cluster file system from any node in the cluster. The exception is ioctls, which you must issue only from the primary node. If you do not know whether an administration command involves ioctls, issue the command from the primary node.

- If a VxFS cluster file system fails over to a secondary node, all standard system-call operations that were in progress during failover are reissued transparently on the new primary. However, any ioctl-related operation in progress during the failover will fail. After a VxFS cluster file-system failover, check the state of the cluster file system. Administrative commands that were issued on the old primary before failover might require corrective measures. See VxFS documentation for more information.

Administering Device Groups

As your cluster requirements change, you might need to add, remove, or modify the device groups on your cluster. Sun Cluster provides an interactive interface called clsetup that you can use to make these changes. clsetup generates cluster commands. Generated commands are shown in the examples at the end of some procedures. The following table lists tasks for administering device groups and provides links to the appropriate procedures in this section.

Note – Sun Cluster software automatically creates a raw-disk device group for each disk and tape device in the cluster. However, cluster device groups remain in an offline state until you access the groups as global devices.

TABLE 5–3 Task Map: Administering Device Groups

Task	Instructions
Update the global device namespace without a reconfiguration reboot by using cldevice populate	"How to Update the Global Device Namespace" on page 137
Add Solaris Volume Manager disk sets and register them as device groups by using metaset	"How to Add and Register a Device Group (Solaris Volume Manager)" on page 138
Remove Solaris Volume Manager device groups from the configuration by using metaset and metaclear	"How to Remove and Unregister a Device Group (Solaris Volume Manager)" on page 140
Remove a node from all device groups by using cldevicegroup, metaset, and clsetup	"How to Remove a Node From All Device Groups" on page 140
Remove a node from a Solaris Volume Manager device group by using metaset	"How to Remove a Node From a Device Group (Solaris Volume Manager)" on page 141
- Use	

TABLE 5–3 Task Map: Administering Device Groups   *(Continued)*

Task	Instructions
SPARC: Add VERITAS Volume Manager disk groups as device groups by using VxVM commands and `clsetup`	"SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 144
	"SPARC: How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 145
	"SPARC: How to Add a New Volume to an Existing Device Group (VERITAS Volume Manager)" on page 146
	"SPARC: How to Convert an Existing Disk Group to a Device Group (VERITAS Volume Manager)" on page 147
	"SPARC: How to Assign a New Minor Number to a Device Group (VERITAS Volume Manager)" on page 148
	"SPARC: How to Register a Disk Group as a Device Group (VERITAS Volume Manager)" on page 149
	"How to Convert a Local Disk Group to a Device Group (VxVM)" on page 152
	"How to Convert a Device Group to a Local Disk Group (VxVM)" on page 153
	"SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 151
SPARC: Remove VERITAS Volume Manager device groups from the configuration by using `clsetup` (to generate `cldevicegroup`)	"SPARC: How to Remove a Volume From a Device Group (VERITAS Volume Manager)" on page 154
	"SPARC: How to Remove and Unregister a Device Group (VERITAS Volume Manager)" on page 155
SPARC: Add a node to a VERITAS Volume Manager device group by using `clsetup` to generate `cldevicegroup`	"SPARC: How to Add a Node to a Device Group (VERITAS Volume Manager)" on page 156
SPARC: Remove a node from a VERITAS Volume Manager device group by using `clsetup` to generate `cldevicegroup`	"SPARC: How to Remove a Node From a Device Group (VERITAS Volume Manager)" on page 157
Remove a node from a raw-disk device group by using `cldevicegroup`	"How to Remove a Node From a Raw-Disk Device Group" on page 159
Change device group properties by using `clsetup` to generate `cldevicegroup`	"How to Change Device Group Properties" on page 161
Display device groups and properties by using `cldevicegroup show`	"How to List a Device Group Configuration" on page 165

TABLE 5–3 Task Map: Administering Device Groups   *(Continued)*

Task	Instructions
Change the desired number of secondaries for a device group by using clsetup to generate cldevicegroup	"How to Set the Desired Number of Secondaries for a Device Group" on page 162
Switch the primary for a device group by using cldevicegroup switch	"How to Switch the Primary for a Device Group" on page 166
Put a device group in maintenance state by using metaset or vxdg	"How to Put a Device Group in Maintenance State" on page 167

▼ How to Update the Global Device Namespace

When adding a new global device, manually update the global device namespace by running the cldevice populate command.

Note – The cldevice populate command does not have any effect if the node that is running the command is not currently a cluster member. The command also has no effect if the /global/.devices/node@ *nodeID* file system is not mounted.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2 **Reconfigure the namespace.**

```
# cldevice populate
```

Example 5–15 Updating the Global Device Namespace

The following example shows the output generated by successfully running the cldevice populate command.

```
# cldevice populate
Configuring the /dev/global directory (global devices)...
obtaining access to all attached disks
reservation program successfully exiting
```

▼ How to Add and Register a Device Group (Solaris Volume Manager)

Use the `metaset` command to create a Solaris Volume Manager disk set and register the disk set as a Sun Cluster device group. When you register the disk set, the name that you assigned to the disk set is automatically assigned to the device group.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on one of the nodes connected to the disks where you are creating the disk set.**

2 **SPARC: Solaris 9 only: Calculate the number of names for Solstice DiskSuite metadevices or Solaris Volume Manager volumes that you need for your configuration, and modify the** `/kernel/drv/md.conf` **file on each node. This step is not required if you are running on Solaris 10.**

See "How to Set the Number of Metadevice or Volume Names and Disk Sets " in *Sun Cluster Software Installation Guide for Solaris OS*.

3 **Use the** `metaset`**(1M) command to add the Solaris Volume Manager disk set and register it as a device group with Sun Cluster. To create a multi-owner disk group, use the** –M **option.**

```
# metaset -s diskset -a -M -h nodelist
```

`-s` *diskset*	Specifies the disk set to be created.
`-a -h` *nodelist*	Adds the list of nodes that can master the disk set.
`-M`	Designates the disk group as multi-owner.

Note – Running the `metaset` command to set up a Solstice DiskSuite/Solaris Volume Manager device group on a cluster results in one secondary by default, regardless of the number of nodes that are included in that device group. You can change the desired number of secondary nodes by using the `clsetup` utility after the device group has been created. Refer to "How to Set the Desired Number of Secondaries for a Device Group" on page 162 for more information about disk failover.

---

4 **If you are configuring a replicated device group, set the replication property for the device group.**

```
# cldevicegroup sync devicegroup
```

5 **Verify that the device group has been added.**

The device group name matches the disk set name that is specified with `metaset`.

```
# cldevicegroup list
```

6   List the DID mappings.

cldevice show | grep Device

- Choose drives that are shared by the cluster nodes that will master or potentially master the disk set.

- Use the full DID device name, which has the form /dev/did/rdsk/d*N*, when you add a drive to a disk set.

In the following example, the entries for DID device /dev/did/rdsk/d3 indicate that the drive is shared by phys-schost-1 and phys-schost-2.

```
=== DID Device Instances ===
DID Device Name:                          /dev/did/rdsk/d1
  Full Device Path:                         phys-schost-1:/dev/rdsk/c0t0d0
DID Device Name:                          /dev/did/rdsk/d2
  Full Device Path:                         phys-schost-1:/dev/rdsk/c0t6d0
DID Device Name:                          /dev/did/rdsk/d3
  Full Device Path:                         phys-schost-1:/dev/rdsk/c1t1d0
  Full Device Path:                         phys-schost-2:/dev/rdsk/c1t1d0
...
```

7   Add the drives to the disk set.

Use the full DID path name.

metaset -s *setname* **-a /dev/did/rdsk/d***N*

-s *setname*   Specifies the disk set name, which is the same as the device group name.

-a   Adds the drive to the disk set.

Note – Do *not* use the lower-level device name (c*N*t*X*d*Y*) when you add a drive to a disk set. Because the lower-level device name is a local name and not unique throughout the cluster, using this name might prevent the metaset from being able to switch over.

---

8   Verify the status of the disk set and drives.

metaset -s *setname*

Example 5–16   Adding a Solaris Volume Manager Device Group

The following example shows the creation of the disk set and device group with the disk drives /dev/did/rdsk/d1 and /dev/did/rdsk/d2 and verifies that the device group has been created.

```
# metaset -s dg-schost-1 -a -h phys-schost-1

# cldevicegroup list
dg-schost-1
```

```
metaset -s dg-schost-1 -a /dev/did/rdsk/d1 /dev/did/rdsk/d2
```

How to Remove and Unregister a Device Group (Solaris Volume Manager)

Device groups are Solaris Volume Manager disk sets that have been registered with Sun Cluster. To remove a Solaris Volume Manager device group, use the `metaclear` and `metaset` commands. These commands remove the device group with the same name and unregister the disk group as a Sun Cluster device group.

Refer to the Solaris Volume Manager documentation for the steps to remove a disk set.

▼ How to Remove a Node From All Device Groups

Use this procedure to remove a cluster node from all device groups that list the node in their lists of potential primaries.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides `solaris.cluster.modify` **RBAC authorization on the node that you are removing as a potential primary of all device groups.**

2 Determine the device group or groups of which the node to be removed is a member.

Look for the node name in the `Device group node list` for each device group.

```
# cldevicegroup list -v
```

3 If any of the device groups identified in Step 2 are of the device group type `SVM`**, perform the steps in "How to Remove a Node From a Device Group (Solaris Volume Manager)" on page 141 for each device group of that type.**

4 If any of the device groups identified in Step 2 are of the device group type `VxVM`**, perform the steps in "SPARC: How to Remove a Node From a Device Group (VERITAS Volume Manager)" on page 157 for each device group of that type.**

5 Determine the raw-device disk groups of which the node to be removed is a member.

```
# cldevicegroup list -v
```

6 If any of the device groups listed in Step 5 are of the device group types `Disk` **or** `Local_Disk`**, perform the steps in "How to Remove a Node From a Raw-Disk Device Group" on page 159 for each of these device groups.**

7   Verify that the node has been removed from the potential primaries list of all device groups.

The command returns nothing if the node is no longer listed as a potential primary of any device group.

```
# cldevicegroup list -v  nodename
```

▼ How to Remove a Node From a Device Group (Solaris Volume Manager)

Use this procedure to remove a cluster node from the list of potential primaries of a Solaris Volume Manager device group. Repeat the metaset command for each device group from which you want to remove the node.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Verify that the node is still a member of the device group and that the device group is a Solaris Volume Manager device group.

Device group type SDS/SVM indicates a Solaris Volume Manager device group.

```
phys-schost-1% cldevicegroup show  devicegroup
```

2   Determine which node is the current primary for the device group.

```
# cluster status -t  devicegroup
```

3   Become superuser on the node that currently owns the device group that you want to modify.

4   Delete the node's host name from the device group.

```
# metaset -s  setname -d -h  nodelist
```

-s *setname*   Specifies the device group name.

-d   Deletes from the device group the nodes identified with -h.

-h *nodelist*   Specifies the nodename of the node or nodes that will be removed.

Note – The update can take several minutes to complete.

If the command fails, add the -f (force) option to the command.

```
# metaset -s  setname -d -f -h  nodelist
```

5   Repeat Step 4 for each device group from which the node is being removed as a potential primary.

6   Verify that the node has been removed from the device group.

The device group name matches the disk set name that is specified with metaset.

```
phys-schost-1% cldevicegroup list -v devicegroup
```

Example 5–17   Removing a Node From a Device Group (Solaris Volume Manager)

The following example shows the removal of the host name phys-schost-2 from a device group configuration. This example eliminates phys-schost-2 as a potential primary for the designated device group. Verify removal of the node by running the cldevicegroup show command. Check that the removed node is no longer displayed in the screen text.

```
[Determine the Solaris Volume Manager
 device group for the node:]
# cldevicegroup show dg-schost-1
=== Device Groups ===

Device Group Name:                      dg-schost-1
  Type:                                 SVM
  failback:                             no
  Node List:                            phys-schost-1, phys-schost-2
  preferenced:                          yes
  numsecondaries:                       1
  diskset name:                         dg-schost-1
[Determine which node is the current primary for the device group:]
# cldevicegroup status dg-schost-1
=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name     Primary          Secondary        Status
----------------      -------          ---------        ------
dg-schost-1           phys-schost-1    phys-schost-2    Online
[Become superuser on the node that currently owns the device group.]
[Remove the host name from the device group:]
# metaset -s dg-schost-1 -d -h phys-schost-2
[Verify removal of the node:]
phys-schost-1% cldevicegroup list -v dg-schost-1
=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name     Primary          Secondary        Status
----------------      -------          ---------        ------
dg-schost-1           phys-schost-1    -                Online
```

▼ How to Create More Than Three Disk Sets in a Cluster

If you are running Solaris 9 and intend to create more than three disk sets in the cluster, perform the following steps before you create the disk sets. You do not need to perform this procedure if you are running Solaris 10. Follow these steps if you are installing disk sets for the first time or if you are adding more disk sets to a fully configured cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Ensure that the value of the** md_nsets **variable is high enough. The value should accommodate the total number of disk sets you intend to create in the cluster.**

    a. On any node of the cluster, check the value of the md_nsets **variable in the** /kernel/drv/md.conf **file.**

    b. If the number of disk sets in the cluster will be greater than the existing value of md_nsets **minus one, increase the value of** md_nsets **on each node.**

    The maximum permissible number of disk sets is the value of md_nsets minus one. The maximum possible value of md_nsets is 32.

    c. Ensure that the /kernel/drv/md.conf **file is identical on each node of the cluster.**

Caution – Failure to follow this guideline can result in serious Solaris Volume Manager errors and possible loss of data.

---

    d. From one node, shut down the cluster.

```
# cluster shutdown -g0 -y
```

    e. Reboot each node in the cluster.

      ■ On SPARC based systems, do the following:

```
ok boot
```

      ■ On x86 based systems, do the following:

      When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                             |
| Solaris failsafe                                                   |
|                                                                   |
```

```
+--------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

2   On each node in the cluster, run the devfsadm**(1M) command.**

You can run this command on all nodes in the cluster at the same time.

3   From one node of the cluster, run the cldevice populate **command.**

4   On each node, verify that the cldevice populate **command has been completed before you attempt to create any disk sets.**

The cldevice command calls itself remotely on all nodes, even when the command is run from just one node. To determine whether the cldevice populate command has completed processing, run the following command on each node of the cluster.

% **ps -ef | grep scgdevs**

▼ SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)

Note – This procedure is only for initializing disks. If you are encapsulating disks, use the procedure "SPARC: How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 145.

After adding the VxVM disk group, you need to register the device group.

If you use VxVM to set up shared disk groups for Oracle Parallel Server or Oracle RAC, use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*.

1   Become superuser on any cluster node that is *physically connected* **to disks that make up the disk group being added.**

2   Create the VxVM disk group and volume.

Use your preferred method to create the disk group and volume.

Note – If you are setting up a mirrored volume, use Dirty Region Logging (DRL) to decrease volume recovery time after a node failure. However, DRL might decrease I/O throughput.

See the VERITAS Volume Manager documentation for the procedures to complete this step.

3 **Register the VxVM disk group as a Sun Cluster device group.**

See "SPARC: How to Register a Disk Group as a Device Group (VERITAS Volume Manager)" on page 149.

Do not register the Oracle Parallel Server or Oracle RAC shared disk groups with the cluster framework.

▼ SPARC: How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)

Note – This procedure is only for encapsulating disks. If you are initializing disks, use the procedure "SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 144.

You can convert nonroot disks to Sun Cluster device groups by encapsulating the disks as VxVM disk groups, then registering the disk groups as Sun Cluster device groups.

Disk encapsulation is only supported during initial creation of a VxVM disk group. After a VxVM disk group is created and registered as a Sun Cluster device group, only disks which can be initialized should be added to the disk group.

If you use VxVM to set up shared disk groups for Oracle Parallel Server or Oracle RAC, use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2 **If the disk being encapsulated has file system entries in the** /etc/vfstab **file, make sure that the** mount at boot **option is set to** no**.**

Set back to yes after the disk is encapsulated and registered as a Sun Cluster device group.

3 **Encapsulate the disks.**

Use vxdiskadm menus or the graphical user interface to encapsulate the disks. VxVM requires two free partitions as well as unassigned cylinders at the beginning or the end of the disk. Slice two must also be set to the entire disk. See the vxdiskadm man page for more information.

4 Shut down and restart the node.

The clnode evacuate command switches over all resource groups and device groups including all non-global zones from the specified node to a next preferred node. Use shutdown to shut down and restart the node.

```
# clnode evacuate   node[,...]
# shutdown -g0 -y -i6
```

5 If necessary, switch all resource groups and device groups back to the original node.

If the resource groups and device groups were initially configured to fail back to the primary node, this step is not necessary.

```
# cldevicegroup switch -n  node devicegroup
# clresourcegroup switch  -z zone  -n  node  resourcegroup
```

node   The name of the node.

zone   The name of the non-global zone on *node* that can master the resource group. Specify *zone* only if you specified a non-global zone when you created the resource group.

6 Register the VxVM disk group as a Sun Cluster device group.

See "SPARC: How to Register a Disk Group as a Device Group (VERITAS Volume Manager)" on page 149.

Do not register the Oracle Parallel Server or Oracle RAC shared disk groups with the cluster framework.

7 If you set the mount at boot **option to** no **in Step 2, set it back to** yes**.**

▼ SPARC: How to Add a New Volume to an Existing Device Group (VERITAS Volume Manager)

When you add a new volume to an existing VxVM device group, perform the procedure from the primary node of the online device group.

Note – After adding the volume, you need to register the configuration change by using the procedure "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 151.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides solaris.cluster.read **and** solaris.cluster.administer **RBAC authorization on any node of the cluster.**

2   Determine the primary node for the device group to which you are adding the new volume.

```
# cldevicegroup status
```

3   If the device group is offline, bring the device group online.

```
# cldevicegroup switch -n nodename devicegroup
```

nodename   Specifies the name of the node to which to switch the device group. This node becomes the new primary.

devicegroup   Specifies the device group to switch.

4   From the primary node (the node currently mastering the device group), create the VxVM volume in the disk group.

Refer to your VERITAS Volume Manager documentation for the procedure used to create the VxVM volume.

5   Synchronize the VxVM disk group changes to update the global namespace.

DPM

▼ SPARC: How to Convert an Existing Disk Group to a Device Group (VERITAS Volume Manager)

You can convert an existing VxVM disk group to a Sun Cluster device group by importing the disk group onto the current node, then registering the disk group as a Sun Cluster device group.

1   Become superuser on any node of the cluster.

2   Import the VxVM disk group to the current node.

```
# vxdg import diskgroup
```

3   Register the VxVM disk group as a Sun Cluster device group.

▼ SPARC: How to Assign a New Minor Number to a Device Group (VERITAS Volume Manager)

If device group registration fails because of a minor number conflict with another disk group, you must assign the new disk group a new, unused minor number. After assigning the new minor number, rerun the procedure to register the disk group as a Sun Cluster device group.

1   Become superuser on any node of the cluster.

2   Determine the minor numbers in use.

```
# ls -l /global/.devices/node@nodeid/dev/vx/dsk/*
```

3   Choose another multiple of 1000 not in use as the base minor number for the new disk group.

4   Assign the new minor number to the disk group.

```
# vxdg reminor diskgroup base-minor-number
```

5   Register the VxVM disk group as a Sun Cluster device group.

See "SPARC: How to Register a Disk Group as a Device Group (VERITAS Volume Manager)" on page 149.

Example 5–18   SPARC: How to Assign a New Minor Number to a Device Group

This example uses the minor numbers 16000-16002 and 4000-4001. The vxdg reminor command is used to assign the base minor number 5000 to the new device group.

```
# ls -l /global/.devices/node@nodeid/dev/vx/dsk/*

/global/.devices/node@nodeid/dev/vx/dsk/dg1
brw-------   1 root     root       56,16000 Oct  7 11:32 dg1v1
brw-------   1 root     root       56,16001 Oct  7 11:32 dg1v2
brw-------   1 root     root       56,16002 Oct  7 11:32 dg1v3

/global/.devices/node@nodeid/dev/vx/dsk/dg2
brw-------   1 root     root       56,4000 Oct  7 11:32 dg2v1
brw-------   1 root     root       56,4001 Oct  7 11:32 dg2v2
# vxdg reminor dg3 5000
```

▼ SPARC: How to Register a Disk Group as a Device Group (VERITAS Volume Manager)

This procedure uses the clsetup utility to register the associated VxVM disk group as a Sun Cluster device group.

Note – After a device group has been registered with the cluster, never import or export a VxVM disk group by using VxVM commands. If you make a change to the VxVM disk group or volume, follow the procedure "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 151 to register the device group configuration changes. This procedure ensures that the global namespace is in the correct state.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

Before You Begin   Ensure that the following prerequisites have been completed prior to registering a VxVM device group:

- Superuser privilege on a node in the cluster.
- The name of the VxVM disk group to be registered as a device group.
- A preferred order of nodes to master the device group.
- A desired number of secondary nodes for the device group.

When you define the preference order, you also specify whether the device group should be switched back to the most preferred node if that node fails and later returns to the cluster.

See cldevicegroup(1CL) for more information about node preference and failback options.

Nonprimary cluster nodes (spares) transition to secondary according to the node preference order. The default number of secondaries for a device group is normally set to one. This default setting minimizes performance degradation that is caused by primary checkpointing of multiple secondary nodes during normal operation. For example, in a four-node cluster, the default behavior configures one primary, one secondary, and two spare nodes. See also "How to Set the Desired Number of Secondaries for a Device Group" on page 162.

1   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2   **Start the** clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

3   To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.

The Device Groups Menu is displayed.

4   To register a VxVM device group, type the number that corresponds to the option for registering a VxVM disk group as a device group.

Follow the instructions and type the name of the VxVM disk group to be registered as a Sun Cluster device group.

If this device group is replicated by using controller-based replication, this name must match the TrueCopy replication group name.

If you use VxVM to set up shared disk groups for Oracle Parallel Server/Oracle RAC, you do not register the shared disk groups with the cluster framework. Use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide.*

5   If you encounter the following error while attempting to register the device group, reminor the device group.

```
scconf: Failed to add device group - in use
```

To reminor the device group, use the procedure "SPARC: How to Assign a New Minor Number to a Device Group (VERITAS Volume Manager)" on page 148. This procedure enables you to assign a new minor number that does not conflict with a minor number that an existing device group uses.

6   If you are configuring a replicated device group, set the replication property for the device group.

cldevicegroup sync *devicegroup*

7   Verify that the device group is registered and online.

If the device group is properly registered, information for the new device group is displayed when you use the following command.

cldevicegroup status *devicegroup*

Note – If you change any configuration information for a VxVM disk group or volume that is registered with the cluster, you must synchronize the device group by using clsetup. Such configuration changes include adding or removing volumes, as well as changing the group, owner, or permissions of existing volumes. Reregistration after configuration changes ensures that the global namespace is in the correct state. See "How to Update the Global Device Namespace" on page 137.

Example 5–19   SPARC: Registering a VERITAS Volume Manager Device Group

The following example shows the cldevicegroup command generated by clsetup when it registers a VxVM device group (dg1), and the verification step. This example assumes that the VxVM disk group and volume were created previously.

```
# clsetup

# cldevicegroup create -t vxvm -n phys-schost-1,phys-schost-2 -p failback=true dg1


# cldevicegroup status dg1

=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name    Primary        Secondary      Status
----------------     -------        ---------      ------
dg1                  phys-schost-1  phys-schost-2  Online
```

See Also   To create a cluster file system on the VxVM device group, see "How to Add a Cluster File System" on page 173.

Is problems occur with the minor number, see "SPARC: How to Assign a New Minor Number to a Device Group (VERITAS Volume Manager)" on page 148.

▼ SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)

When you change any configuration information for a VxVM disk group or volume, you need to register the configuration changes for the Sun Cluster device group. Registration ensures that the global namespace is in the correct state.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on any node in the cluster.**

2   **Start the** clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

3   **To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

4 **To register configuration changes, type the number that corresponds to the option for synchronizing volume information for a VxVM device group.**

Follow the instructions and type the name of the VxVM disk group that has changed configuration.

Example 5–20 SPARC: Registering VERITAS Volume Manager Disk Group Configuration Changes

The following example shows the cldevicegroup command generated by clsetup a changed VxVM device group (dg1) is registered. This example assumes that the VxVM disk group and volume were created previously.

```
# clsetup

cldevicegroup sync dg1
```

▼ How to Convert a Local Disk Group to a Device Group (VxVM)

Perform this procedure to change a local VxVM disk group to a globally accessible VxVM device group.

1 **Become superuser on a node of the cluster.**

2 **Start the** clsetup **utility**

```
# clsetup
```

3 **Unset the** localonly **property.**

a. **Choose the menu item, Device groups and volumes.**

b. **Choose the menu item, Reset a local VxVM disk group to a VxVM device group.**

c. **Follow the instructions to unset the** localonly **property.**

4 **Specify the nodes that can master the disk group.**

a. **Return to the main menu in the** clsetup **utility.**

b. **Choose the menu item, Device groups and volumes.**

c. **Choose the menu item, Register a diskgroup.**

d. **Follow the instructions to specify the nodes that can master the disk group.**

e. **When finished, quit the** `clsetup` **utility.**

5 **Verify that the device group is configured.**

    `phys-schost# ` **`cldevicegroup show`**

▼ How to Convert a Device Group to a Local Disk Group (VxVM)

Perform this procedure to change a VxVM device group to a local VxVM disk group that is not managed by Sun Cluster software. The local disk group can have more than one node in its node list, but it can be mastered by only one node at a time.

1 **Become superuser on a node of the cluster.**

2 **Take the device group offline.**

    `phys-schost# ` **`cldevicegroup offline`** *devicegroup*

3 **Unregister the device group.**

    a. **Start the** `clsetup` **utility.**

        `phys-schost# ` **`clsetup`**

    b. **Choose the menu item, Device groups and volumes.**

    c. **Choose the menu item, Unregister a VxVM disk group.**

    d. **Follow the instructions to specify the VxVM disk group that you are unregistering from Sun Cluster software.**

    e. **Quit the** `clsetup` **utility.**

4 **Verify that the disk group is no longer registered with Sun Cluster software.**

    `phys-schost# ` **`cldevicegroup status`**

    Command output should no longer show the device group that you unregistered.

5 **Import the disk group.**

    `phys-schost# ` **`vxdg import`** *diskgroup*

6   **Set the** localonly **property of the disk group.**

a. **Start the** clsetup **utility.**

phys-schost# **clsetup**

b. **Choose the menu item, Device groups and volumes.**

c. **Choose the menu item, Set a VxVM disk group as a local disk group.**

d. **Follow the instructions to set the** localonly **property and to specify the single node that is to exclusively master the disk group.**

e. **When finished, quit the** clsetup **utility.**

7   **Verify that the disk group is successfully configured as a local disk group.**

phys-schost# **vxdg list** *diskgroup*

▼ SPARC: How to Remove a Volume From a Device Group (VERITAS Volume Manager)

Note – After removing the volume from the device group, you must register the configuration changes to the device group by using the procedure "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 151.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.read **and** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2   **Determine the primary node and status for the device group.**

cldevicegroup status *devicegroup*

3   **If the device group is offline, bring it online.**

cldevicegroup online *devicegroup*

4   **From the primary node (the node currently mastering the device group), remove the VxVM volume in the disk group.**

vxedit -g *diskgroup* **-rf rm** *volume*

| -g *diskgroup* | Specifies the VxVM disk group that contains the volume. |
| -rf rm *volume* | Removes the specified volume. The -r option makes the operation recursive. The -f option is required to remove an enabled volume. |

5   Using the clsetup **utility, register the device group configuration changes to update the global namespace.**

See "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 151.

▼ SPARC: How to Remove and Unregister a Device Group (VERITAS Volume Manager)

Removing a Sun Cluster device group causes the corresponding VxVM disk group to be exported, not destroyed. However, even though the VxVM disk group still exists, it cannot be used in the cluster unless reregistered.

This procedure uses the clsetup utility to remove a VxVM disk group and unregister it as a Sun Cluster device group.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2   Take the device group offline.

    # **cldevicegroup offline** *devicegroup*

3   Start the clsetup **utility.**

    # **clsetup**

The Main Menu is displayed.

4   To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.

The Device Groups Menu is displayed.

5   To unregister a VxVM disk group, type the number that corresponds to the option for unregistering a VxVM device group.

Follow the instructions and type the name of the VxVM disk group to be unregistered.

Example 5–21   SPARC: Removing and Unregistering a VERITAS Volume Manager Device Group

The following example shows the VxVM device group dg1 taken offline, and the cldevicegroup command generated by clsetup when it removes and unregisters the device group.

```
# cldevicegroup offline dg1
# clsetup

  cldevicegroup delete dg1
```

▼ SPARC: How to Add a Node to a Device Group (VERITAS Volume Manager)

This procedure adds a node to a device group using the clsetup utility.

The prerequisites to add a node to a VxVM device group are:

- Superuser privilege on a node in the cluster
- The name of the VxVM device group to which the node will be added
- The name or node ID of the nodes to add

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.read **and** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2   **Start the** clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

3   **To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

4   **To add a node to a VxVM device group, type the number that corresponds to the option for adding a node to a VxVM device group.**

Follow the instructions and type the device group and node names.

5   **Verify that the node has been added.**

Look for the device group information for the new disk displayed by the following command.

```
# cldevicegroup show devicegroup
```

Example 5–22 SPARC: Adding a Node to a VERITAS Volume Manager Device Group

The following example shows the scconf command generated by clsetup when it adds a node (phys-schost-3) to a VxVM device group (dg1), and the verification step.

```
# clsetup

cldevicegroup add-node -n phys-schost-3 dg1

# cldevicegroup show dg1

=== Device Groups ===

Device Group Name:                    dg1
  Type:                               VxVM
  failback:                           yes
  Node List:                          phys-schost-1, phys-schost-3
  preferenced:                        no
  numsecondaries:                     1
  diskgroup names:                       dg1
```

▼ SPARC: How to Remove a Node From a Device Group (VERITAS Volume Manager)

Use this procedure to remove a cluster node from the list of potential primaries of a VERITAS Volume Manager (VxVM) device group (disk group).

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Verify that the node is still a member of the group and that the group is an VxVM device group.**

Device group type VxVM indicates a VxVM device group.

```
phys-schost-1% cldevicegroup show devicegroup
```

2 **Become superuser or assume a role that provides** solaris.cluster.read **and** solaris.cluster.modify **RBAC authorization on a current cluster member node.**

3 **Start the** clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

4 To reconfigure a device group, type the number that corresponds to the option for device groups and volumes.

5 To remove the node from the VxVM device group, type the number that corresponds to the option for removing a node from a VxVM device group.

Follow the prompts to remove the cluster node from the device group. You are asked for information about the following:

- VxVM device group
- Node name

6 Verify that the node has been removed from the VxVM device group or groups.

```
# cldevicegroup show devicegroup
```

Example 5–23 SPARC: Removing a Node From a Device Group (VxVM)

This example shows removal of the node named phys-schost-1 from the dg1 VxVM device group.

[*Determine the VxVM device group for the node:*]
```
# cldevicegroup show dg1


=== Device Groups ===

Device Group Name:                        dg1
  Type:                                   VXVM
  failback:                               no
  Node List:                              phys-schost-1, phys-schost-2
  preferenced:                            no
  numsecondaries:                         1
  diskgroup names:                            dg1
```
[*Become superuser and start the* clsetup *utility*:]
```
# clsetup
```
  Select Device groups and volumes>Remove a node from a VxVM device group.

  Answer the questions when prompted.
```
You will need the following information.
  Name:        Example:
  VxVM device group name    dg1
  node names               phys-schost-1
```

[*Verify that the cldevicegroup command executed properly:*]
```
 cldevicegroup remove-node -n phys-schost-1 dg1


   Command completed successfully.
```
Dismiss the clsetup   **Device Groups Menu and Main Menu.**
[*Verify that the node was removed:*]

```
# cldevicegroup show dg1

=== Device Groups ===

Device Group Name:                      dg1
  Type:                                 VXVM
  failback:                             no
  Node List:                            phys-schost-2
  preferenced:                          no
  numsecondaries:                       1
  device names:                         dg1
```

▼ How to Remove a Node From a Raw-Disk Device Group

Use this procedure to remove a cluster node from the list of potential primaries of a raw-disk device group.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.read and solaris.cluster.modify RBAC authorization on a node in the cluster *other than the node to remove*.

2   Identify the device groups that are connected to the node being removed, and determine which are raw-disk device groups.

    # **cldevicegroup show -n** *nodename* **-t rawdisk +**

3   Disable the localonly property of each Local_Disk raw-disk device group.

    # **cldevicegroup set -p localonly=false** *devicegroup*

    See the cldevicegroup(1CL) man page for more information about the localonly property.

4   Verify that you have disabled the localonly property of all raw-disk device groups that are connected to the node being removed.

    The Disk device group type indicates that the localonly property is disabled for that raw-disk device group.

    # **cldevicegroup show -n** *nodename* **-t rawdisk -v +**

5   Remove the node from all raw-disk device groups that are identified in Step 2.

    You must complete this step for each raw-disk device group that is connected to the node being removed.

    # **cldevicegroup remove-node -n** *nodename devicegroup*

Example 5–24   SPARC: Removing a Node From a Raw Device Group

This example shows how to remove a node (phys-schost-2) from a raw-disk device group. All commands are run from another node of the cluster (phys-schost-1).

[*Identify the device groups connected to the node being removed, and determine which are raw-disk device groups:*]
```
phys-schost-1# cldevicegroup show -n phys-schost-2 -t rawdisk -v +
Device Group Name:                          dsk/d4
  Type:                                     Disk
  failback:                                 false
  Node List:                                phys-schost-2
  preferenced:                              false
  localonly:                                false
  autogen                                   true
  numsecondaries:                           1
  device names:                             phys-schost-2

Device Group Name:                          dsk/d2
  Type:                                     VxVM
  failback:                                 true
  Node List:                                pbrave2
  preferenced:                              false
  localonly:                                false
  autogen                                   true
  numsecondaries:                           1
  diskgroup name:                           vxdg1

Device Group Name:                          dsk/d1
  Type:                                     SVM
  failback:                                 false
  Node List:                                pbrave1, pbrave2
  preferenced:                              true
  localonly:                                false
  autogen                                   true
  numsecondaries:                           1
  diskset name:                             ms1
(dsk/d4) Device group node list:  phys-schost-2
    (dsk/d2) Device group node list:  phys-schost-1, phys-schost-2
    (dsk/d1) Device group node list:  phys-schost-1, phys-schost-2
```
[*Disable the localonly flag for each local disk on the node:*]
```
phys-schost-1# cldevicegroup set -p localonly=false dsk/d4
```
[*Verify that the localonly flag is disabled:*]
```
phys-schost-1# cldevicegroup show -n phys-schost-2 -t rawdisk +
 (dsk/d4) Device group type:          Disk
 (dsk/d8) Device group type:          Local_Disk
```
[*Remove the node from all raw-disk device groups:*]

```
phys-schost-1# cldevicegroup remove-node -n phys-schost-2 dsk/d4
phys-schost-1# cldevicegroup remove-node -n phys-schost-2 dsk/d2
phys-schost-1# cldevicegroup remove-node -n phys-schost-2 dsk/d1
```

▼ How to Change Device Group Properties

The method for establishing the primary ownership of a device group is based on the setting of an ownership preference attribute called preferenced. If the attribute is not set, the primary owner of an otherwise unowned device group is the first node that attempts to access a disk in that group. However, if this attribute is set, you must specify the preferred order in which nodes attempt to establish ownership.

If you disable the preferenced attribute, then the failback attribute is also automatically disabled. However, if you attempt to enable or re-enable the preferenced attribute, you have the choice of enabling or disabling the failback attribute.

If the preferenced attribute is either enabled or re-enabled, you are required to re-establish the order of nodes in the primary ownership preference list.

This procedure uses clsetup to set or unset the preferenced attribute and the failback attribute for Solaris Volume Manager or VxVM device groups.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

Before You Begin   To perform this procedure, you need the name of the device group for which you are changing attribute values.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.read **and** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2   **Start the** clsetup **utility.**

clsetup

The Main Menu is displayed.

3   **To work with device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

4 **To change key properties of a device group, type the number that corresponds to the option for changing key properties of a VxVM or Solaris Volume Manager device group).**

The Change Key Properties Menu is displayed.

5 **To change a device group property, type the number that corresponds to the option for changing the preferenced and/or failback properties.**

Follow the instructions to set the preferenced and failback options for a device group.

6 **Verify that the device group attributes have been changed.**

Look for the device group information displayed by the following command.

```
# cldevicegroup show -v devicegroup
```

Example 5–25 Changing Device Group Properties

The following example shows the cldevicegroup command generated by clsetup when it sets the attribute values for a device group (dg-schost-1).

```
# cldevicegroup set -p preferenced=true -p failback=true -p numsecondaries=1 \
-p nodelist=phys-schost-1,phys-schost-2 dg-schost-1
# cldevicegroup show dg-schost-1

=== Device Groups ===

Device Group Name:                    dg-schost-1
  Type:                                SVM
  failback:                            yes
  Node List:                           phys-schost-1, phys-schost-2
  preferenced:                         yes
  numsecondaries:                      1
  diskset names:                        dg-schost-1
```

▼ How to Set the Desired Number of Secondaries for a Device Group

The numsecondaries property specifies the number of nodes within a device group that can master the group if the primary node fails. The default number of secondaries for device services is one. You can set the value to any integer between one and the number of operational nonprimary provider nodes in the device group.

This setting is an important factor in balancing cluster performance and availability. For example, increasing the desired number of secondaries increases the device group's opportunity to survive multiple failures that occur simultaneously within a cluster. Increasing the number of secondaries also decreases performance regularly during normal operation. A smaller number of secondaries

typically results in better performance, but reduces availability. However, a larger number of secondaries does not always result in greater availability of the file system or device group in question. Refer to Chapter 3, "Key Concepts for System Administrators and Application Developers," in *Sun Cluster Concepts Guide for Solaris OS* for more information.

If you change the numsecondaries property, secondary nodes are added or removed from the device group if the change causes a mismatch between the actual number of secondaries and the desired number.

This procedure uses the clsetup utility to set the numsecondaries property for all types of device groups. Refer to cldevicegroup(1CL) for information about device group options when configuring any device group.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.read **and** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2   **Start the** clsetup **utility.**

clsetup

The Main Menu is displayed.

3   **To work with device groups, select the option labeled Device groups and volumes.**

The Device Groups Menu is displayed.

4   **To change key properties of a device group, select the option labeled Change key properties of a device group.**

The Change Key Properties Menu is displayed.

5   **To change the desired number of secondaries, type the number that corresponds to the option for changing the** numsecondaries **property.**

Follow the instructions and type the desired number of secondaries to be configured for the device group. The corresponding cldevicegroup command is then executed, a log is printed, and the utility returns to the previous menu.

6   **Validate the device group configuration.**

cldevicegroup show dg-schost-1
=== Device Groups ===

Device Group Name: dg-schost-1
  Type: VxVm   *This might also be SDS or Local_Disk.*
  failback: yes

```
Node List:                              phys-schost-1, phys-schost-2 phys-schost-3
preferenced:                            yes
numsecondaries:                         1
diskgroup names:                            dg-schost-1
```

Note – If you change any configuration information for a VxVM disk group or volume that is registered with the cluster, you must reregister the device group by using `clsetup`. Such configuration changes include adding or removing volumes, as well as changing the group, owner, or permissions of existing volumes. Reregistration after configuration changes ensures that the global namespace is in the correct state. See "How to Update the Global Device Namespace" on page 137.

7 Verify that the device group attribute has been changed.

Look for the device group information that is displayed by the following command.

```
# cldevicegroup show -v devicegroup
```

Example 5–26 Changing the Desired Number of Secondaries (Solstice DiskSuite or Solaris Volume Manager)

The following example shows the `cldevicegroup` command that is generated by `clsetup` when it configures the desired number of secondaries for a device group (`dg-schost-1`). This example assumes that the disk group and volume were created previously.

```
# cldevicegroup set -p numsecondaries=1 dg-schost-1
# cldevicegroup show -v dg-schost-1

=== Device Groups ===

Device Group Name:                      dg-schost-1
  Type:                                 SVM
  failback:                             yes
  Node List:                            phys-schost-1, phys-schost-2
  preferenced:                          yes
  numsecondaries:                       1
  diskset names:                          dg-schost-1
```

Example 5–27 SPARC: Setting the Desired Number of Secondaries (VERITAS Volume Manager)

The following example shows the `cldevicegroup` command that is generated by `clsetup` when it sets the desired number of secondaries for a device group (`dg-schost-1`) to two. See "How to Set the Desired Number of Secondaries for a Device Group" on page 162 for information about changing the desired number of secondaries after a device group is created.

```
# cldevicegroup set -p numsecondaries=2 dg-schost-1
```

```
# cldevicegroup show dg-schost-1
=== Device Groups ===

Device Group Name:                   dg-schost-1
  Type:                              VxVM
  failback:                          yes
  Node List:                         phys-schost-1, phys-schost-2
  preferenced:                       yes
  numsecondaries:                    1
  diskgroup names:                      dg-schost-1
```

Example 5–28 Setting the Desired Number of Secondaries to the Default Value

The following example shows use of a null string value to configure the default number of secondaries. The device group will be configured to use the default value, even if the default value changes.

```
# cldevicegroup set -p numsecondaries= dg-schost-1
# cldevicegroup show -v dg-schost-1

=== Device Groups ===

Device Group Name:                   dg-schost-1
  Type:                              SVM
  failback:                          yes
  Node List:                         phys-schost-1, phys-schost-2 phys-schost-3
  preferenced:                       yes
  numsecondaries:                    1
  diskset names:                       dg-schost-1
```

▼ How to List a Device Group Configuration

You do not need to be superuser to list the configuration. However, you do need solaris.cluster.read authorization.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

◗ **Use one method from the following list.**

Sun Cluster Manger GUI	See the Sun Cluster Manager online help for more information.
cldevicegroup show	Use cldevicegroup show to list the configuration for all device groups in the cluster.

cldevicegroup show *devicegroup*   Use cldevicegroup show *devicegroup* to list the
                                          configuration of a single device group.

cldevicegroup status *devicegroup*   Use cldevicegroup status *devicegroup* to determine the
                                          status of a single device group.

cldevicegroup status +   Use cldevicegroup status + to determine the status of
                                          all device groups in the cluster.

Use the -v option with any of these commands to obtain more detailed information.

Example 5–29   Listing the Status of All Device Groups

```
# cldevicegroup status +

=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name    Primary         Secondary       Status
----------------     -------         ---------       ------
dg-schost-1          phys-schost-2   phys-schost-1   Online
dg-schost-2          phys-schost-1   --              Offline
dg-schost-3          phys-schost-3   phy-shost-2     Online
```

Example 5–30   Listing the Configuration of a Particular Device Group

```
# cldevicegroup show dg-schost-1

=== Device Groups ===

Device Group Name:                          dg-schost-1
  Type:                                       SVM
  failback:                                   yes
  Node List:                                  phys-schost-2, phys-schost-3
  preferenced:                                yes
  numsecondaries:                             1
  diskset names:                               dg-schost-1
```

▼ How to Switch the Primary for a Device Group

This procedure can also be used to start (bring online) an inactive device group.

You can also bring an inactive device group online or switch the primary for a device group by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a profile that provides** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2 **Use** cldevicegroup switch **to switch the device group primary.**

cldevicegroup switch -n *nodename devicegroup*

-n *nodename*   Specifies the name of the node to switch to. This node become the new primary.

devicegroup   Specifies the device group to switch.

3 **Verify that the device group has been switched to the new primary.**

If the device group is properly registered, information for the new device group is displayed when you use the following command.

cldevice status *devicegroup*

Example 5–31   Switching the Primary for a Device Group

The following example shows how to switch the primary for a device group and verify the change.

```
# cldevicegroup switch -n phys-schost-1 dg-schost-1

# cldevicegroup status dg-schost-1

=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name     Primary         Secondary       Status
----------------     -------         ---------       ------
dg-schost-1          phys-schost-1   phys-schost-2   Online
```

▼ **How to Put a Device Group in Maintenance State**

Putting a device group in maintenance state prevents that device group from automatically being brought online whenever one of its devices is accessed. You should put a device group in maintenance state when completing repair procedures that require that all I/O activity be acquiesced until completion of the repair. Putting a device group in maintenance state also helps prevent data loss by ensuring that a device group is not brought online on one node while the disk set or disk group is being repaired on another node.

> **Note –** Before a device group can be placed in maintenance state, all access to its devices must be stopped, and all dependent file systems must be unmounted.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1. **Place the device group in maintenance state.**

   a. **If the device group is enabled, disable the device group.**

      ```
      # cldevicegroup disable devicegroup
      ```

   b. **Take the device group offline.**

      ```
      # cldevicegroup offline devicegroup
      ```

2. **If the repair procedure being performed requires ownership of a disk set or disk group, manually import that disk set or disk group.**

   For Solaris Volume Manager:

   ```
   # metaset -C take -f -s diskset
   ```

> **Caution –** If you are taking ownership of a Solaris Volume Manager disk set, you *must* use the `metaset -C take` command when the device group is in maintenance state. Using `metaset -t` brings the device group online as part of taking ownership. If you are importing a VxVM disk group, you must use the `-t` flag when importing the disk group. Using the `-t` flag prevents the disk group from automatically being imported if this node is rebooted.

   For VERITAS Volume Manager:

   ```
   # vxdg -t import disk-group-name
   ```

3. **Complete the repair procedure that you need to perform.**

4. **Release ownership of the disk set or disk group.**

> **Caution –** Before taking the device group out of maintenance state, you must release ownership of the disk set or disk group. Failure to release ownership can result in data loss.

   - For Solaris Volume Manager:

     ```
     # metaset -C release -s diskset
     ```

   - For VERITAS Volume Manager:

```
# vxdg deport diskgroupname
```

5 Bring the device group online.

```
# cldevicegroup online devicegroup
# cldevicegroup enable devicegroup
```

Example 5–32 Putting a Device Group in Maintenance State

This example shows how to put device group dg-schost-1 in maintenance state, and remove the device group from maintenance state.

[*Place the device group in maintenance state.*]
```
# cldevicegroup disable dg-schost-1
# cldevicegroup offline dg-schost-1
```
[*If needed, manually import the disk set or disk group.*]
For Solaris Volume Manager:
```
  # metaset -C take -f -s dg-schost-1
```
For VERITAS Volume Manager:
```
  # vxdg -t import dg1
```

[*Complete all necessary repair procedures.*]

[*Release ownership.*]
For Solaris Volume Manager:
```
  # metaset -C release -s dg-schost-1
```
For VERITAS Volume Manager:
```
  # vxdg deport dg1
```

[*Bring the device group online.*]
```
# cldevicegroup online dg-schost-1
# cldevicegroup enable dg-schost-1
```

Administering the SCSI Protocol Settings for Storage Devices

Sun Cluster software installation automatically assigns SCSI reservations to all storage devices. Use the following procedures to check the settings of devices and, if necessary, to override the setting for a device.

- "How to Display the Default Global SCSI Protocol Settings for All Storage Devices" on page 170
- "How to Display the SCSI Protocol of a Single Storage Device" on page 170
- "How to Change the Default Global SCSI Protocol Settings for All Storage Devices " on page 171
- "How to Change the SCSI Protocol for a Single Storage Device" on page 172

▼ How to Display the Default Global SCSI Protocol Settings for All Storage Devices

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.read **RBAC authorization.**

2 **From any node, display the current global default SCSI protocol setting.**

```
# cluster show -t global
```

For more information, see the cluster(1CL) man page.

Example 5–33 Displaying the Default Global SCSI Protocol Settings for All Storage Devices

The following example displays the SCSI protocol settings for all storage devices on the cluster.

```
# cluster show -t global

=== Cluster ===

Cluster Name:                           racerxx
  installmode:                          disabled
  heartbeat_timeout:                    10000
  heartbeat_quantum:                    1000
  private_netaddr:                      172.16.0.0
  private_netmask:                      255.255.248.0
  max_nodes:                            64
  max_privatenets:                      10
  global_fencing:                       scsi3
  Node List:                            phys-racerxx-1, phys-racerxx-2
```

▼ How to Display the SCSI Protocol of a Single Storage Device

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.read **RBAC authorization.**

2 From any node, display the SCSI protocol setting of the storage device.

cldevice show *device*

device   The name of the device path or a device name.

For more information, see the cldevice(1CL)man page.

Example 5–34   Displaying the SCSI Protocol of a Single Device

The following example displays the SCSI protocol for the device /dev/rdsk/c4t8d0.

```
# cldevice show /dev/rdsk/c4t8d0


=== DID Device Instances ===

DID Device Name:                              /dev/did/rdsk/d3
  Full Device Path:                             phappy1:/dev/rdsk/c4t8d0
  Full Device Path:                             phappy2:/dev/rdsk/c4t8d0
  Replication:                                  none
  default_fencing:                              global
```

▼ How to Change the Default Global SCSI Protocol Settings for All Storage Devices

The default fencing setting of a single storage device overrides the global setting when the device's default fencing is set to pathcount or prefer3. If the default fencing setting of a storage device is set to global, the storage device will use the global setting. For example, if a storage device has the default setting pathcount, the setting will not change if you use this procedure to change the global SCSI protocol settings to prefer3. You must use the "How to Change the SCSI Protocol for a Single Storage Device" on page 172 procedure to change the default setting of a single device.

To change the default fencing setting for a quorum device you must unconfigure the device, change the setting, and reconfigure the quorum device.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization.**

2 Set the SCSI protocol for all storage devices that are not quorum devices.

cluster set -p global_fencing={prefer3 | pathcount}

-p global_fencing   Sets the current global default fencing algorithm for all shared devices.

prefer3   Uses the SCSI-3 protocol.

pathcount   Determines the fencing protocol by the number of DID paths that are
                             attached to the shared device.

Example 5–35   Setting the Default Global SCSI Protocol Settings for All Storage Devices

The following example sets the SCSI protocol for all storage devices on the cluster to the SCSI-3
protocol.

```
# cluster set -p global_fencing=prefer3
```

▼ How to Change the SCSI Protocol for a Single Storage Device

To change the default fencing setting for a quorum device you must unconfigure the device.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have
short forms. Except for the long and short forms of the command names, the commands are
identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization.**

2   Set the SCSI protocol of the storage device.

```
# cldevice set -p default_fencing ={pathcount | prefer3 | global} device
```

-p   Modifies the device properties.

pathcount   Determines the fencing protocol by the number of DID paths that are attached to the
              shared device.

prefer3   Uses the SCSI-3 protocol.

global   Uses the global default fencing setting.

device   Specifies the name of the device path or device name.

For more information, see the cluster(1CL) man page.

Example 5–36   Setting the SCSI Protocol of a Single Device

The following example sets the device 11, specified by device number, to the SCSI-3 protocol.

```
# cldevice set -p default_fencing=prefer3 11
```

Administering Cluster File Systems

The cluster file system is a globally available file system that can be read and accessed from any node of the cluster.

TABLE 5–4 Task Map: Administering Cluster File Systems

Task	Instructions
Add cluster file systems after the initial Sun Cluster installation by using newfs(1M) and mkdir	"How to Add a Cluster File System" on page 173
Remove a cluster file system by using fuser(1M) and umount(1M)	"How to Remove a Cluster File System" on page 176
Check global mount points in a cluster for consistency across nodes by using sccheck(1M)	"How to Check Global Mounts in a Cluster" on page 178

▼ How to Add a Cluster File System

Perform this task for each cluster file system you create after your initial Sun Cluster installation.

Caution – Be sure you specify the correct disk device name. Creating a cluster file system destroys any data on the disks. If you specify the wrong device name, you will erase data that you might not intend to delete.

Ensure the following prerequisites have been completed prior to adding an additional cluster file system:

- Superuser privilege is established on a node in the cluster.
- Volume manager software be installed and configured on the cluster.
- A device group (Solaris Volume Manager device group or a VxVM device group) or block disk slice exists on which to create the cluster file system.

If you used Sun Cluster Manager to install data services, one or more cluster file systems already exist if shared disks on which to create the cluster file systems were sufficient.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser on any node in the cluster.

Tip – For faster file system creation, become superuser on the current primary of the global device for which you are creating a file system.

---

2   Create a file system by using the newfs **command.**

Note – The newfs command is valid only for creating new UFS file systems. To create a new VxFS file system, follow the procedures provided in your VxFS documentation.

---

newfs *raw-disk-device*

The following table shows examples of names for the *raw-disk-device* argument. Note that naming conventions differ for each volume manager.

Volume Manager	Disk Device Name	Description
Solaris Volume Manager	/dev/md/oracle/rdsk/d1	Raw disk device d1 within the oracle disk set.
SPARC: VERITAS Volume Manager	/dev/vx/rdsk/oradg/vol01	Raw disk device vol01 within the oradg disk group.
None	/dev/global/rdsk/d1s3	Raw disk device for block slice d1s3.

3   On each node in the cluster, create a mount-point directory for the cluster file system.

A mount point is required *on each node*, even if the cluster file system will not be accessed on that node.

Tip – For ease of administration, create the mount point in the /global/*devicegroup* directory. Using this location enables you to easily distinguish cluster file systems, which are globally available, from local file systems.

---

mkdir -p /global/devicegroup mountpoint

devicegroup   Name of the directory that corresponds to the name of the device group that contains the device.

mountpoint   Name of the directory on which to mount the cluster file system.

4   On each node in the cluster, add an entry to the /etc/vfstab **file for the mount point.**

a.   Use the following required mount options.

Note – Logging is required for all cluster file systems.

■ **Solaris UFS logging** – Use the global, logging mount options. See the mount_ufs(1M) man page for more information about UFS mount options.

Note – The syncdir mount option is not required for UFS cluster file systems. If you specify syncdir, you are guaranteed POSIX-compliant file-system behavior. If you do not, you will experience the same behavior as with UFS file systems. When you do not specify syncdir, performance of writes that allocate disk blocks, such as when appending data to a file, can significantly improve. However, in some cases, without syncdir you would not discover an out-of-space condition until you close a file. The cases in which you could have problems if you do not specify syncdir are rare. With syncdir (and POSIX behavior), the out-of-space condition would be discovered before the close.

■ **Solaris Volume Manager transactional volume** – Use the global mount option (do not use the logging mount option). See your Solaris Volume Manager documentation for information about setting up transactional volumes.

Note – Transactional volumes are scheduled to be removed from the Solaris OS in an upcoming Solaris software release. Solaris UFS logging provides the same capabilities but superior performance, as well as lower system administration requirements and overhead.

■ **VxFS logging** – Use the global and log mount options. See the mount_vxfs man page that is provided with VxFS software for more information.

b. **To automatically mount the cluster file system, set the** mount at boot **field to** yes**.**

c. **Ensure that, for each cluster file system, the information in its** /etc/vfstab **entry is identical on each node.**

d. **Ensure that the entries in each node's** /etc/vfstab **file list devices in the same order.**

e. **Check the boot order dependencies of the file systems.**

For example, consider the scenario where phys-schost-1 mounts disk device d0 on /global/oracle, and phys-schost-2 mounts disk device d1 on /global/oracle/logs. With this configuration, phys-schost-2 can boot and mount /global/oracle/logs only after phys-schost-1 boots and mounts /global/oracle .

See the vfstab(4) man page for details.

5 On any node in the cluster, verify that mount points exist and /etc/vfstab **file entries are correct on all nodes of the cluster.**

sccheck

If no errors occur, nothing is returned.

6   From any node in the cluster, mount the cluster file system.

    # **mount /global/***devicegroup mountpoint*

7   On each node of the cluster, verify that the cluster file system is mounted.

You can use either the df or mount command to list mounted file systems.

To manage a VxFS cluster file system in a Sun Cluster environment, run administrative commands only from the primary node on which the VxFS cluster file system is mounted.

Example 5–37   Adding a Cluster File System

The following example creates a UFS cluster file system on the Solaris Volume Manager metadevice or volume /dev/md/oracle/rdsk/d1.

```
# newfs /dev/md/oracle/rdsk/d1
...

[on each node:]
# mkdir -p /global/oracle/d1

# vi /etc/vfstab
#device               device               mount          FS fsck  mount   mount
#to mount             to fsck              point          type pass at boot options
# /dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes global,logging

[save and exit]

[on one node:]
# sccheck
# mount /dev/md/oracle/dsk/d1 /global/oracle/d1
# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
on Sun Oct 3 08:56:16 2001
```

▼   **How to Remove a Cluster File System**

You *remove* a cluster file system by merely unmounting it. To also remove or delete the data, remove the underlying disk device (or metadevice or volume) from the system.

> **Note –** Cluster file systems are automatically unmounted as part of the system shutdown that occurs when you run `cluster shutdown` to stop the entire cluster. A cluster file system is not unmounted when you run `shutdown` to stop a single node. However, if the node being shut down is the only node with a connection to the disk, any attempt to access the cluster file system on that disk results in an error.

Ensure that the following prerequisites have been completed prior to unmounting cluster file systems:

- Superuser privilege is established on a node in the cluster.
- The file system is not busy. A file system is considered busy if a user is working in a directory in the file system, or if a program has a file open in that file system. The user or program could be running on any node in the cluster.

1   Become superuser on any node in the cluster.

2   Determine which cluster file systems are mounted.

```
# mount -v
```

3   On each node, list all processes that are using the cluster file system, so that you know which processes you are going to stop.

```
# fuser -c [ -u ] mountpoint
```

-c   Reports on files that are mount points for file systems and any files within those mounted file systems.

-u   (Optional) Displays the user login name for each process ID.

mountpoint   Specifies the name of the cluster file system for which you want to stop processes.

4   On each node, stop all processes for the cluster file system.

Use your preferred method for stopping processes. If necessary, use the following command to force termination of processes associated with the cluster file system.

```
# fuser -c -k mountpoint
```

A SIGKILL is sent to each process that uses the cluster file system.

5   On each node, verify that no processes are using the file system.

```
# fuser -c mountpoint
```

6   From just one node, unmount the file system.

```
# umount mountpoint
```

mountpoint   Specifies the name of the cluster file system you want to unmount. This can be either the directory name where the cluster file system is mounted, or the device name path of the file system.

7 (Optional) Edit the /etc/vfstab **file to delete the entry for the cluster file system being removed.**

Perform this step on each cluster node that has an entry for this cluster file system in its /etc/vfstab file.

8 (Optional) Remove the disk device group/metadevice/volume/plex**.**

See your volume manager documentation for more information.

Example 5–38   Removing a Cluster File System

The following example removes a UFS cluster file system that is mounted on the Solaris Volume Manager metadevice or volume/dev/md/oracle/rdsk/d1.

```
# mount -v
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
# fuser -c /global/oracle/d1
/global/oracle/d1: 4006c
# fuser -c -k /global/oracle/d1
/global/oracle/d1: 4006c
# fuser -c /global/oracle/d1
/global/oracle/d1:
# umount /global/oracle/d1
```

(*On each node, remove the highlighted entry:*)
```
# vi /etc/vfstab
#device         device         mount  FS     fsck   mount   mount
#to mount       to fsck        point  type   pass   at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes global,logging
```

[*Save and exit.*]

To remove the data on the cluster file system, remove the underlying device. See your volume manager documentation for more information.

▼ **How to Check Global Mounts in a Cluster**

The sccheck(1M) utility verifies the syntax of the entries for cluster file systems in the /etc/vfstab file. If no errors occur, nothing is returned.

> **Note –** Run sccheck after making cluster configuration changes, such as removing a cluster file
> system, that have affected devices or volume management components.

1 Become superuser on any node in the cluster.

2 Check the cluster global mounts.

```
# sccheck
```

Administering Disk-Path Monitoring

Disk path monitoring (DPM) administration commands enable you to receive notification of
secondary disk-path failure. Use the procedures in this section to perform administrative tasks that
are associated with monitoring disk paths. Refer to Chapter 3, "Key Concepts for System
Administrators and Application Developers," in *Sun Cluster Concepts Guide for Solaris OS* for
conceptual information about the disk-path monitoring daemon. Refer to the cldevice(1CL) man
page for a description of the scdpm command options and related commands. Refer to the
syslogd(1M) man page for logged errors that the daemon reports.

> **Note –** Disk paths are automatically added to the monitoring list monitored when I/O devices are
> added to a node by using the cldevice command. Disk paths are also automatically unmonitored
> when devices are removed from a node by using Sun Cluster commands.

TABLE 5–5 Task Map: Administering Disk-Path Monitoring

Task	Instructions
Monitor a disk path by using the cldevice monitor command	"How to Monitor a Disk Path" on page 180
Unmonitor a disk path by using the cldevice unmonitor command	"How to Unmonitor a Disk Path" on page 181
Print the status of faulted disk paths for a node by using the cldevice status command	"How to Print Failed Disk Paths" on page 182
Monitor disk paths from a file by using the cldevice command	"How to Monitor Disk Paths From a File" on page 182
Enable or disable the automatic rebooting of a node when all monitored disk paths fail by using the clnode set command	"How to Enable the Automatic Rebooting of a Node When All Monitored Disk Paths Fail" on page 184
	"How to Disable the Automatic Rebooting of a Node When All Monitored Disk Paths Fail" on page 185

The procedures in the following section that issue the `cldevice` command include the disk-path argument. The disk-path argument consists of a node name and a disk name. The node name is not required and defaults to `all` if you do not specify it.

▼ How to Monitor a Disk Path

Perform this task to monitor disk paths in your cluster.

⚠️ **Caution** – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on any node in the cluster.**

2 **Monitor a disk path.**

```
# cldevice monitor -n node disk
```

3 **Verify that the disk path is monitored.**

```
# cldevice status device
```

Example 5–39 Monitoring a Disk Path on a Single Node

The following example monitors the `schost-1:/dev/did/rdsk/d1` disk path from a single node. Only the DPM daemon on the node `schost-1` monitors the path to the disk `/dev/did/dsk/d1`.

```
# cldevice monitor -n schost-1 /dev/did/dsk/d1
# cldevice status d1

Device Instance    Node          Status
--------------- ---- ------
/dev/did/rdsk/d1   phys-schost-1 Ok
```

Example 5–40 Monitoring a Disk Path on All Nodes

The following example monitors the `schost-1:/dev/did/dsk/d1` disk path from all nodes. DPM starts on all nodes for which `/dev/did/dsk/d1` is a valid path.

```
# cldevice monitor /dev/did/dsk/d1
# cldevice status /dev/did/dsk/d1
```

```
Device Instance    Node         Status
--------------- ---- ------
/dev/did/rdsk/d1   phys-schost-1 Ok
```

Example 5–41   Rereading the Disk Configuration From the CCR

The following example forces the daemon to reread the disk configuration from the CCR and prints the monitored disk paths with status.

```
# cldevice monitor +
# cldevice status
Device Instance               Node            Status
---------------               ----            ------
/dev/did/rdsk/d1              schost-1        Ok
/dev/did/rdsk/d2              schost-1        Ok
/dev/did/rdsk/d3              schost-1        Ok
                               schost-2        Ok
/dev/did/rdsk/d4              schost-1        Ok
                               schost-2        Ok
/dev/did/rdsk/d5              schost-1        Ok
                               schost-2        Ok
/dev/did/rdsk/d6              schost-1        Ok
                               schost-2        Ok
/dev/did/rdsk/d7              schost-2        Ok
/dev/did/rdsk/d8              schost-2        Ok
```

▼ How to Unmonitor a Disk Path

Use this procedure to unmonitor a disk path.

> ⚠ **Caution –** DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on any node in the cluster.**

2   Determine the state of the disk path to unmonitor.

```
# cldevice status device
```

3 **On each node, unmonitor the appropriate disk paths.**

```
# cldevice unmonitor -n node disk
```

Example 5–42 Unmonitoring a Disk Path

The following example unmonitors the `schost-2:/dev/did/rdsk/d1` disk path and prints disk paths with status for the entire cluster.

```
# cldevice unmonitor -n schost2 /dev/did/rdsk/d1
# cldevice status -n schost2 /dev/did/rdsk/d1

Device Instance          Node            Status
---------------          ----            ------
/dev/did/rdsk/d1         schost-2        Unmonitored
```

▼ How to Print Failed Disk Paths

Use the following procedure to print the faulted disk paths for a cluster.

⚠️ **Caution** – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

1 **Become superuser on any node in the cluster.**

2 **Print the faulted disk paths throughout the cluster.**

```
# cldevice status -s fail
```

Example 5–43 Printing Faulted Disk Paths

The following example prints faulted disk paths for the entire cluster.

```
# cldevice status -s fail

Device Instance          Node            Status
---------------          ----            ------
dev/did/dsk/d4           phys-schost-1   fail
```

▼ How to Monitor Disk Paths From a File

Use the following procedure to monitor or unmonitor disk paths from a file.

To change your cluster configuration by using a file, you must first export the current configuration. This operation creates an XML file that you can then modify to set the configuration items you are changing. The instructions in this procedure describe this entire process.

Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on any node in the cluster.**

2   **Export your device configuration to an XML file.**

    # **cldevice export -o** *configurationfile*

    -o *configurationfile*   Specify the file name for your XML file.

3   **Modify the configuration file so that device paths are monitored.**

    Find the device paths that you want to monitor, and set the monitored attribute to true.

4   **Monitor the device paths.**

    # **cldevice monitor -i** *configurationfile*

    -i *configurationfile*   Specify the file name of the modified XML file.

5   **Verify that device path is now monitored.**

    # **cldevice status**

Example 5–44   Monitor Disk Paths From a File

In the following example, the device path between the node phys-schost–2 and device d3 is monitored by using an XML file.

The first step is to export the current cluster configuration.

    # **cldevice export -o deviceconfig**

The deviceconfig XML file shows that the path between phys-schost–2 and d3 is not currently monitored.

```
<?xml version="1.0"?>
<!DOCTYPE cluster SYSTEM "/usr/cluster/lib/xml/cluster.dtd">
```

```
<cluster name="brave_clus">
.
.
.
  <deviceList readonly="true">
   <device name="d3" ctd="c1t8d0">
     <devicePath nodeRef="phys-schost-1" monitored="true"/>
     <devicePath nodeRef="phys-schost-2" monitored="false"/>
   </device>
  </deviceList>
</cluster>
```

To monitor that path, set the monitored attribute to true, as follows.

```
<?xml version="1.0"?>
<!DOCTYPE cluster SYSTEM "/usr/cluster/lib/xml/cluster.dtd">
<cluster name="brave_clus">
.
.
.
  <deviceList readonly="true">
   <device name="d3" ctd="c1t8d0">
     <devicePath nodeRef="phys-schost-1" monitored="true"/>
     <devicePath nodeRef="phys-schost-2" monitored="true"/>
   </device>
  </deviceList>
</cluster>
```

Use the cldevice command to read the file and turn on monitoring.

```
# cldevice monitor -i deviceconfig
```

Use the cldevice command to verify that the device is now monitored.

```
# cldevice status
```

See Also   For more details about exporting cluster configuration and using the resulting XML file to set cluster configuration, see the cluster(1CL) and the clconfiguration(5CL) man pages.

▼ How to Enable the Automatic Rebooting of a Node When All Monitored Disk Paths Fail

When you enable this feature, a node automatically reboots, provided that the following conditions are met:

- All monitored disk paths on the node fail.
- At least one of the disks is accessible from a different node in the cluster.

Rebooting the node restarts all resource groups and device groups that are mastered on that node on another node.

If all monitored disk paths on a node remain inaccessible after the node automatically reboots, the node does not automatically reboot again. However, if any disk paths become available after the node reboots but then fail, the node automatically reboots again.

1 **On any node in the cluster, become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization.**

2 **For** *all* **nodes in the cluster, enable the automatic rebooting of a node when all monitored disk paths to it fail.**

```
# clnode set -p reboot_on_path_failure=enabled +
```

▼ How to Disable the Automatic Rebooting of a Node When All Monitored Disk Paths Fail

When you disable this feature and all monitored disk paths on a node fail, the node does *not* automatically reboot.

1 **On any node in the cluster, become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization.**

2 **For** *all* **nodes in the cluster, disable the automatic rebooting of a node when monitored all monitored disk paths to it fail.**

```
# clnode set -p reboot_on_path_failure=disabled +
```

6

Administering Quorum

This chapter provides the procedures for administering quorum within Sun Cluster. For information about quorum concepts, see "Quorum and Quorum Devices" in *Sun Cluster Concepts Guide for Solaris OS*.

Most examples shown in this chapter are from a three-node cluster.

Overview of Administering Quorum

You can use the clquorum(1CL) command to perform all quorum administrative procedures. In addition, you can accomplish some procedures by using the clsetup(1CL) interactive utility or the Sun Cluster Manager GUI. Whenever possible, quorum procedures are described in this chapter by using the clsetup utility. The Sun Cluster Manager online help describes how to perform quorum procedures by using the GUI. When you work with quorum devices, keep in mind the following guidelines:

- If you are running the Solaris 10 OS, then all quorum commands must be run in the global zone.
- If the clquorum command is interrupted or fails, the quorum configuration information can become inconsistent in the cluster configuration database. If this inconsistency occurs, either rerun the command or run the clquorum reset command to reset the quorum configuration.
- For highest availability of the cluster, ensure that the total number of votes that are contributed by quorum devices is less than the total number of votes that are contributed by nodes. Otherwise, the nodes cannot form a cluster if all quorum devices are unavailable, even if all nodes are functioning.

- Ensure that the spanning tree algorithm is disabled on any router or switch between the quorum server and the cluster nodes. An enabled spanning tree algorithm might cause the router or switch to not permit communications for extended periods of time. This prevention of communication to the quorum server is eventually interpreted by the cluster as loss of the quorum device.

- Do not add a disk that is currently configured as a quorum device to a Solaris ZFS storage pool. When a configured quorum device is added to a Solaris ZFS storage pool, the disk is relabeled as an EFI disk and quorum configuration information is lost and the disk no longer provides a quorum vote to the cluster. Once a disk is in a storage pool, that disk can then be configured as a quorum device. Or, you can unconfigure the disk, add it to the storage pool, then reconfigure the disk as a quorum device.

Note – The clsetup command is an interactive interface to the other Sun Cluster commands. When clsetup runs, the command generates the appropriate specific commands, in this case clquorum commands. These generated commands are shown in the examples at the end of the procedures.

To view the quorum configuration, use clquorum show. The clquorum list command displays the names of quorum devices in the cluster. The clquorum status command provides status and vote count information.

TABLE 6–1 Task List: Administering Quorum

Task	For Instructions
Add a quorum device to a cluster by using clsetup(1CL)	"Adding a Quorum Device" on page 190
Remove a quorum device from a cluster by using clsetup (to generate clquorum)	"How to Remove a Quorum Device" on page 198
Remove the last quorum device from a cluster by using clsetup (to generate clquorum)	"How to Remove the Last Quorum Device From a Cluster" on page 199
Replace a quorum device in a cluster by using the add and remove procedures	"How to Replace a Quorum Device" on page 200
Modify a quorum device list by using the add and remove procedures	"How to Modify a Quorum Device Node List" on page 201
Put a quorum device into maintenance state by using clsetup (to generate clquorum) (While in maintenance state, the quorum device does not participate in voting to establish the quorum.)	"How to Put a Quorum Device Into Maintenance State" on page 204

TABLE 6–1 Task List: Administering Quorum   *(Continued)*

Task	For Instructions
Reset the quorum configuration to its default state by using `clsetup` (to generate clquorum)	"How to Bring a Quorum Device Out of Maintenance State" on page 205
List the quorum devices and vote counts by using `clquorum(1CL)`	"How to List the Quorum Configuration" on page 207

Dynamic Reconfiguration With Quorum Devices

You must consider a few issues when completing dynamic reconfiguration (DR) operations on quorum devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support, except for the operating system quiescence operation. Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect nonnetwork IO devices during a DR detach operation.

- Sun Cluster rejects DR remove-board operations that are performed when an interface is present that is configured for a quorum device.

- If the DR operation would pertain to an active device, Sun Cluster rejects the operation and identifies the devices that would be affected by the operation.

To remove a quorum device, you must complete the following steps, in the order indicated.

TABLE 6–2 Task Map: Dynamic Reconfiguration With Quorum Devices

Task	For Instructions
1. Enable a new quorum device to replace the one being removed.	"Adding a Quorum Device" on page 190
2. Disable the quorum device to be removed.	"How to Remove a Quorum Device" on page 198
3. Perform the DR remove operation on the device being removed.	*Sun Enterprise 10000 DR Configuration Guide* and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (from the *Solaris 9 on Sun Hardware*, and *Solaris 10 on Sun Hardware* collections.)

Adding a Quorum Device

This section provides procedures to add a quorum device. For information about determining the number of quorum vote counts necessary for your cluster, recommended quorum configurations, and failure fencing, see "Quorum and Quorum Devices" in *Sun Cluster Concepts Guide for Solaris OS*.

> ⚠️ **Caution** – Do not add a disk that is currently configured as a quorum device to a Solaris ZFS storage pool. When a configured quorum device is added to a Solaris ZFS storage pool, the disk is relabeled as an EFI disk and quorum configuration information is lost and the disk no longer provides a quorum vote to the cluster. Once a disk is in a storage pool, that disk can then be configured as a quorum device. Or, you can unconfigure the disk, add it to the storage pool, then reconfigure the disk as a quorum device.

Sun Cluster supports the following types of quorum devices: SCSI, Network Appliance (NetApp) NAS, and Sun Cluster Quorum Server. Procedures for adding these devices are provided in the following sections:

- "How to Add a SCSI Quorum Device" on page 190
- "How to Add a Network Appliance Network-Attached Storage (NAS) Quorum Device" on page 192
- "How to Add a Quorum Server Quorum Device" on page 195

> **Note** – You cannot configure replicated disks as quorum devices. If you attempt to add a replicated disk as a quorum device, you will receive the following error message, after which the command exits with an error code.
>
> ```
> Disk-name is a replicated device. Replicated devices cannot be
> configured as quorum devices.
> ```

You can also accomplish these procedures by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

See the `clsetup`(1CL) and `clquorum`(1CL) man pages for information about the commands used in the following procedures.

▼ How to Add a SCSI Quorum Device

To complete this procedure, identify a disk drive by its device ID (DID), which is shared by the nodes. Use the `cldevice show` command to see the list of DID names. Refer to the `cldevice`(1CL) man page for additional information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on any node of the cluster.**

2   **Start the** `clsetup` **utility.**

    # **clsetup**

    The `clsetup` Main Menu is displayed.

3   **Type the number that corresponds to the option for Quorum.**

    The Quorum Menu is displayed.

4   **Type the number that corresponds to the option for adding a quorum device, then type yes when the** `clsetup` **utility asks you to confirm the quorum device that you are adding.**

    The `clsetup` utility asks what type of quorum device you want to add.

5   **Type the number that corresponds to the option for a SCSI quorum device.**

    The `clsetup` utility asks which global device you want to use.

6   **Type the global device you are using.**

    The `clsetup` utility asks you to confirm that the new quorum device should be added to the global device you specified.

7   **Type yes to continue adding the new quorum device.**

    If the new quorum device is added successfully, the `clsetup` utility displays a message to that effect.

8   **Verify that the quorum device has been added.**

    # **clquorum list -v**

Example 6–1   Adding a SCSI Quorum Device

    The following example shows the `clquorum` command generated by `clsetup` when it adds a SCSI quorum device and a verification step.

Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.

```
[Start the clsetup utility:]
# clsetup
[Select Quorum>Add a quorum device]
[Answer the questions when prompted.]
[You will need the following information.]
   [Information:        Example:]
   [SCSI device         scsi]
   [Global device       d20]

[Verify that the clquorum command was completed successfully:]
```

```
 clquorum add d20

   Command completed successfully.
```
[*Quit the clsetup Quorum Menu and Main Menu.*]
[*Verify that the quorum device is added:*]
 # clquorum list -v

```
Quorums         Type
-------         ----
d20             scsi
scphyshost-1    node
scphyshost-2    node
```

▼ How to Add a Network Appliance Network-Attached Storage (NAS) Quorum Device

When you use a Network Appliance (NetApp) network-attached storage (NAS) device as a quorum device, the following are required:

- You must install the iSCSI license from NetApp.

- You must configure an iSCSI LUN on the clustered filer for use as the quorum device.

- You must configure the NetApp NAS unit to use NTP for synchronizing time.

- At least one of the NTP servers selected for the clustered filer must be an NTP server for the Sun Cluster nodes.

- When booting the cluster, always boot the NAS device before you boot the cluster nodes.

  If you boot devices in the wrong order, your nodes cannot find the quorum device. If a node should fail in this situation, your cluster might be unable to remain in service. If a service disruption occurs, you must either reboot the entire cluster or remove the NetApp NAS quorum device and add it again.

- A cluster can use a NAS device for only a single quorum device.

  You can configure other shared storage if you need additional quorum devices. Additional clusters that use the same NAS device can use separate LUNs on that device as their quorum devices.

See the following Network Appliance NAS documentation for information about creating and setting up a Network Appliance NAS device and LUN. You can access the following documents at http://now.netapp.com.

Task	Network Appliance Documentation
Setting up a NAS device	*System Administration File Access Management Guide*

Task	Network Appliance Documentation
Setting up a LUN	*Host Cluster Tool for Unix Installation Guide*
Installing ONTAP software	*Software Setup Guide*, *Upgrade Guide*
Exporting volumes for the cluster	*Data ONTAP Storage Management Guide*
Installing NAS support software packages on cluster nodes	Log in to http://now.netapp.com. From the Software Download page, download the *Host Cluster Tool for Unix Installation Guide*.

See the following Sun Cluster documentation for information about installing a NetApp NAS storage device in a Sun Cluster environment: *Sun Cluster 3.1 - 3.2 With Network-Attached Storage Devices Manual for Solaris OS.*

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Make sure that all Sun Cluster nodes are online and can communicate with the NetApp clustered filer.

2   Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on any node of the cluster.**

3   Start the clsetup **utility.**

clsetup

The clsetup Main Menu is displayed.

4   Type the number that corresponds to the option for Quorum.

The Quorum Menu is displayed.

5   Type the number that corresponds to the option for adding a quorum device. Then type yes to confirm that you are adding a quorum device.

The clsetup utility asks what type of quorum device you want to add.

6   Type the number that corresponds to the option for a netapp_nas quorum device. Then type yes to confirm that you are adding a netapp_nas quorum device.

The clsetup utility asks you provide the name of the new quorum device.

7   Type the name of the quorum device you are adding.

The quorum device name can be any name you choose. The name is only used to process future administrative commands.

The clsetup utility asks you to provide the name of the filer for the new quorum device.

8   Type the name of the filer of the new quorum device.

This name is the network-accessible name or address of the filer.

The clsetup utility asks you to provide the LUN ID for the filer.

9   Type the ID of the quorum device LUN on the filer.

The clsetup utility asks if to the new quorum device should be added on the filer.

10   Type yes to continue adding the new quorum device.

If the new quorum device is added successfully, the clsetup utility displays a message to that effect.

11   Verify that the quorum device has been added.

```
# clquorum list -v
```

Example 6–2   Adding a NetApp NAS Quorum Device

The following example shows the clquorum command generated by clsetup when it adds a NetApp NAS quorum device. The example also shows a verification step.

Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.

[*Start the clsetup utility:*]
```
# clsetup
```
[*Select Quorum>Add a quorum device*]
[*Answer the questions when prompted.*]
[*You will need the following information.*]
  [*Information:*   *Example:*]
  [*Quorum Device*   *Netapp_nas quorum device*]
  [*Name:*   *qd1*]
  [*Filer:*   *nas1.sun.com*]
  [*LUN ID:*   *0*]

[*Verify that the clquorum command was completed successfully:*]
```
 clquorum add -t netapp_nas -p filer=nas1.sun.com,-p lun_id=0 qd1


    Command completed successfully.
```
[*Quit the clsetup Quorum Menu and Main Menu.*]
[*Verify that the quorum device is added:*]
```
# clquorum list -v

Quorums         Type
-------         ----
qd1             netapp_nas
scphyshost-1    node
scphyshost-2    node
```

▼ How to Add a Quorum Server Quorum Device

Before You Begin   Before you can add a Sun Cluster Quorum Server as a quorum device, the Sun Cluster Quorum Server software must be installed on the host machine and the quorum server must be started and running. For information about installing and starting the quorum server, see the *Sun Cluster Quorum Server User's Guide*.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2 **Make sure that all Sun Cluster nodes are online and can communicate with the Sun Cluster Quorum Server.**

   a. On each node in the cluster, add the quorum server host name to the /etc/inet/hosts **file or the** /etc/inet/ipnodes **file.**

   Add a host name-to-address mapping to the file, such as the following.

   ipaddress qshost1

   ipaddress   The IP address of the computer where the quorum server is running.

   qshost1   The host name of the computer where the quorum server is running.

   b. On each cluster node, add the quorum server host's netmask to /etc/inet/netmasks **and add the cluster nodes' netmasks to** /etc/inet/netmasks**.**

   Add an entry to the file that identifies the subnet where the quorum server host is running, such as the following.

   192.168.11.0 255.255.255.0

   192.168.11.0   The IP address of the computer where the quorum server is running.

   255.255.255.0   The netmask of the quorum server host.

3 **Start the** clsetup **utility.**

   # **clsetup**

   The clsetup Main Menu is displayed.

4 **Type the number that corresponds to the option for Quorum.**

   The Quorum Menu is displayed.

5 **Type the number that corresponds to the option for adding a quorum device. Then type yes to confirm that you are adding a quorum device.**

The clsetup utility asks what type of quorum device you want to add.

6 **Type the number that corresponds to the option for a quorum_server quorum device. Then type yes to confirm that you are adding a quorum_server quorum device.**

The clsetup utility asks you to provide the name of the new quorum device.

7 **Type the name of the quorum device you are adding.**

The quorum device name can be any name you choose. The name is only used to process future administrative commands.

The clsetup utility asks you to provide the name of the filer for the new quorum device.

8 **Type the name of the host of the quorum server.**

This name specifies the IP address of the machine where the quorum server runs or the host name of the machine on the network.

Depending on the IPv4 or IPv6 configuration of the host, the IP address of the machine must be specified in the /etc/hosts file, the /etc/inet/ipnodes file, or both.

Note – The machine you specify must be reachable by all cluster nodes and must run the quorum server.

The clsetup utility asks you to provide the port number of the quorum server.

9 **Type the port number that is used by the quorum server to communicate with the cluster nodes.**

The clsetup utility asks you to confirm that the new quorum device should be added.

10 **Type yes to continue adding the new quorum device.**

If the new quorum device is added successfully, the clsetup utility displays a message to that effect.

11 **Verify that the quorum device has been added.**

```
# clquorum list -v
```

Example 6–3 Adding a Quorum Server Quorum Device

The following example shows the clquorum command that is generated by clsetup when it adds a quorum server quorum device. The example also shows a verification step.

Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.

[*Start the clsetup utility:*]
```
# clsetup
```

[*Select Quorum>Add a quorum device*]
[*Answer the questions when prompted.*]
[*You will need the following information.*]
   [*Information: Example:*]
   [*Quorum Device   quorum_server quorum device*]
   [*Name: qd1*]
   [*Host Machine Name: 10.11.124.84*]
   [*Port Number: 9001*]

[*Verify that the clquorum command was completed successfully:*]
 clquorum add -t quorum_server -p qshost=10.11.124.84,-p port=9001 qd1

    Command completed successfully.
[*Quit the clsetup Quorum Menu and Main Menu.*]
[*Verify that the quorum device is added:*]
clquorum list -v

```
Quorums         Type
-------         ----
qd1             quorum_server
scphyshost-1    node
scphyshost-2    node

-- Quorum Summary --

  Quorum votes possible: 5
  Quorum votes needed:   3
  Quorum votes present:  5

-- Quorum Votes by Node --

                   Node Name        Present Possible Status
                   ---------        ------- -------- ------
  Node votes:      phys-schost-1    1       1        Online
  Node votes:      phys-schost-2    1       1        Online

-- Quorum Votes by Device --

                   Device Name      Present Possible Status
                   -----------      ------- -------- ------
  Device votes:    qd1              1       1        Online
  Device votes:    /dev/did/rdsk/d3s2  1    1        Online
  Device votes:    /dev/did/rdsk/d4s2  1    1        Online
```

Removing or Replacing a Quorum Device

This section provides the following procedures for removing or replacing a quorum device:

- "How to Remove a Quorum Device" on page 198
- "How to Remove the Last Quorum Device From a Cluster" on page 199
- "How to Replace a Quorum Device" on page 200

▼ How to Remove a Quorum Device

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

When a quorum device is removed, it no longer participates in the voting to establish quorum. Note that all two-node clusters require that at least one quorum device be configured. If this is the last quorum device on a cluster, clquorum(1CL) will fail to remove the device from the configuration.

Note – If the device you intend to remove is the last quorum device in the cluster, see the procedure "How to Remove the Last Quorum Device From a Cluster" on page 199.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on any node in the cluster.**

2   **Determine the quorum device to be removed.**

     `# clquorum list -v`

3   **Execute the** clsetup**(1CL) utility.**

     `# clsetup`

     The Main Menu is displayed.

4   **Type the number that corresponds to the option for Quorum.**

5   **Type the number that corresponds to the option to remove a quorum device.**
     Answer the questions displayed during the removal process.

6   **Quit** clsetup**.**

7   Verify that the quorum device is removed.

```
# clquorum list -v
```

Example 6–4   Removing a Quorum Device

This example shows how to remove a quorum device from a cluster with two or more quorum devices configured.

Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.

[*Determine the quorum device to be removed:*]
`# clquorum list -v`
[*Start the clsetup utility:*]
`# clsetup`
[*Select Quorum>Remove a quorum device*]
[*Answer the questions when prompted.*]
[*Quit the clsetup Quorum Menu and Main Menu.*]
[*Verify that the quorum device is removed:*]
`# clquorum list -v`

```
 Quorums        Type
 -------        ----
scphyshost-1    node
scphyshost-2    node
scphyshost-3    node
```

Troubleshooting   If you lose communications between the cluster and the quorum server host while removing a quorum server quorum device, you must clean up stale configuration information on the quorum server host. For instructions on performing this cleanup, see the *Sun Cluster Quorum Server User's Guide*.

▼ How to Remove the Last Quorum Device From a Cluster

If the device you intend to remove is not the last quorum device in the cluster, use the previous procedure, "How to Remove a Quorum Device" on page 198.

Note – All two-node clusters require at least one configured quorum device. If this is the last quorum device on a two-node cluster, the cluster must be placed into install mode before clquorum(1CL) enables you to remove the device from the configuration. This procedure should only be performed if a node is being removed from the cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on any node in the cluster and place the node to be removed in maintenance state.**

    See "How to Put a Node Into Maintenance State" on page 239.

2   **Place the cluster in install mode.**

    ```
    # cluster set -p installmode=enabled
    ```

3   **Remove the quorum device by using the** `clquorum` **command.**

    The `clsetup`(1CL) cluster-administration menu options are not available while the cluster is in install mode.

    ```
    # clquorum remove qd1
    ```

4   **Verify that the quorum device has been removed.**

    ```
    # clquorum list -v
    ```

Example 6–5   Removing the Last Quorum Device

This example shows how to remove the last remaining quorum device in a cluster configuration.

[*Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.*]
[*Place the cluster in install mode:*]
```
# cluster set -p installmode=enabled
```
[*Remove the quorum device:*]
```
# clquorum remove d3
```
[*Verify that the quorum device has been removed:*]
```
# clquorum list -v
 Quorums         Type
-------         ----
scphyshost-1    node
scphyshost-2    node
scphyshost-3    node
```

▼ How to Replace a Quorum Device

Use this procedure to replace an existing quorum device with another quorum device. You can replace a quorum device with a similar device type, such as replacing a NAS device with another NAS device, or you can replace the device with a dissimilar device, such as replacing a NAS device with a shared disk.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Configure a new quorum device.

You need to first add a new quorum device to the configuration to take the place of the old device. See "Adding a Quorum Device" on page 190 to add a new quorum device to the cluster.

2 Remove the device that you are replacing as a quorum device.

See "How to Remove a Quorum Device" on page 198 to remove the old quorum device from the configuration.

3 If the quorum device is a failed disk, replace the disk.

Refer to the hardware procedures for your disk enclosure in the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS.*

Maintaining Quorum Devices

This section provides the following procedures for maintaining quorum devices:

- "How to Modify a Quorum Device Node List" on page 201
- "How to Put a Quorum Device Into Maintenance State" on page 204
- "How to Bring a Quorum Device Out of Maintenance State" on page 205
- "How to List the Quorum Configuration" on page 207
- "How to Repair a Quorum Device" on page 208

▼ How to Modify a Quorum Device Node List

You can use the clsetup(1CL) utility to add a node to or remove a node from the node list of an existing quorum device. To modify a quorum device's node list, you must remove the quorum device, modify the physical connections of nodes to the quorum device you removed, then add the quorum device to the cluster configuration again. When a quorum device is added, clquorum(1CL) automatically configures the node-to-disk paths for all nodes attached to the disk.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on any node of the cluster.**

2 Determine the name of the quorum device you are modifying.

```
# clquorum list -v
```

3 Start the clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

4 Type the number that corresponds to the Quorum option.

The Quorum Menu is displayed.

5 Type the number that corresponds to the option to remove a quorum device.

Follow the instructions. You will be asked the name of the disk to be removed.

6 Add or delete the physical node connections to the quorum device.

7 Type the number that corresponds to the option to add a quorum device.

Follow the instructions. You will be asked the name of the disk to be used as the quorum device.

8 Verify that the quorum device has been added.

```
# clquorum list -v
```

Example 6–6 Modifying a Quorum Device Node List

The following example shows how to use the clsetup utility to add nodes to or delete nodes from a quorum device node list. In this example, the quorum device name is d2, and the final result of the procedures adds another node to the quorum device node list.

[*Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any node in the cluster.*]

[*Determine the quorum device name:*]
```
# clquorum list -v
Quorums            Type
-------            -----
d2                 scsi
sc-phys-schost-1   node
sc-phys-schost-2   node
sc-phys-schost-3   node
```

[*Start the clsetup utility:*]
```
# clsetup
```

[*Type the number that corresponds with the quorum option.*]
.
[*Type the number that corresponds with the option to remove a quorum device.*]

.
[*Answer the questions when prompted.*]
[*You will need the following information:*]

   Information: Example:
   Quorum Device Name: d2

[*Verify that the clquorum command completed successfully:*]
 clquorum remove d2
    Command completed successfully.

[*Type the number that corresponds with the Quorum option.*]
.
[*Type the number that corresponds with the option to add a quorum device.*]
.
[*Answer the questions when prompted.*]
[*You will need the following information:*]

   Information   Example:
   quorum device name   d2

[*Verify that the clquorum command was completed successfully:*]
clquorum add d2
    Command completed successfully.

Quit the clsetup utility.

[*Verify that the correct nodes have paths to the quorum device.*
In this example, note that phys-schost-3 has been added to the
enabled hosts list.]
 # clquorum show d2 | grep Hosts
=== Quorum Devices ===

Quorum Device Name: d2
  Hosts (enabled):   phys-schost-1, phys-schost-2, phys-schost-3

[*Verify that the modified quorum device is online.*]
 # clquorum status d2
=== Cluster Quorum ===

--- Quorum Votes by Device ---

```
Device Name        Present      Possible      Status
-----------        -------      --------      ------
d2                 1            1             Online
```
[*Verify the quorum device is removed:*]
 # clquorum list -v
Quorums   Type

```
-------           -----
sc-phys-schost-1   node
sc-phys-schost-2   node
sc-phys-schost-3   node
```

▼ How to Put a Quorum Device Into Maintenance State

Use the clquorum(1CL) command to put a quorum device into maintenance state. The clsetup(1CL) utility does not currently have this capability. You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

Put a quorum device into maintenance state when taking the quorum device out of service for an extended period of time. This way, the quorum device's quorum vote count is set to zero and does not contribute to the quorum count while the device is being serviced. While in maintenance state, the quorum device's configuration information is preserved.

Note – All two-node clusters require at least one configured quorum device. If this is the last quorum device on a two-node cluster, clquorum will fail to put the device into maintenance state.

To put a cluster node into maintenance state, see .

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any node of the cluster.

2   Put the quorum device into the maintenance state.

    # **clquorum disable** *device*

    device   Specifies the DID name of the disk device to change, for example, d4.

3   Verify that the quorum device is now in maintenance state.
    The output for the device you placed in maintenance state should read zero for the Quorum Device Votes.

    # **clquorum status** *device*

Example 6–7 Putting a Quorum Device Into Maintenance State

The following example shows how to put a quorum device into maintenance state and how to verify the results.

```
# clquorum disable d20
# clquorum status d20

=== Cluster Quorum ===

--- Quorum Votes by Device ---

Device Name      Present      Possible      Status
-----------      -------      --------      ------
d20              1            1             Offline
```

See Also To re-enable the quorum device, see "How to Bring a Quorum Device Out of Maintenance State" on page 205.

To put a node into maintenance state, see "How to Put a Node Into Maintenance State" on page 239.

▼ How to Bring a Quorum Device Out of Maintenance State

Run this procedure each time a quorum device is in maintenance state and you want to bring the quorum device out of maintenance state and reset the quorum vote count to the default.

> ⚠ **Caution –** If you do not specify either the globaldev or node options, the quorum count is reset for the entire cluster.

When you configure a quorum device, Sun Cluster software assigns the quorum device a vote count of N-1 where N is the number of connected votes to the quorum device. For example, a quorum device that is connected to two nodes with nonzero vote counts has a quorum count of one (two minus one).

- To bring a cluster node as well as its associated quorum devices out of maintenance state, see "How to Remove Node From Maintenance State" on page 240.
- To learn more about quorum vote counts, see "About Quorum Vote Counts" in *Sun Cluster Concepts Guide for Solaris OS*.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on any node of the cluster.**

2 **Reset the quorum count.**

clquorum enable *device*

device   Specifies the DID name of the quorum device to reset, for example, d4.

3 **If you are resetting the quorum count because a node was in maintenance state, reboot the node.**

4 **Verify the quorum vote count.**

clquorum show +

Example 6–8   Resetting the Quorum Vote Count (Quorum Device)

The following example resets the quorum count for a quorum device back to the default and verifies the result.

```
# clquorum enable d20
# clquorum show +

=== Cluster Nodes ===

Node Name:                               phys-schost-2
  Node ID:                               1
  Quorum Vote Count:                     1
  Reservation Key:                       0x43BAC41300000001

Node Name:                               phys-schost-3
  Node ID:                               2
  Quorum Vote Count:                     1
  Reservation Key:                       0x43BAC41300000002


=== Quorum Devices ===

Quorum Device Name:                      d3
  Enabled:                               yes
  Votes:                                 1
  Global Name:                           /dev/did/rdsk/d20s2
  Type:                                  scsi
  Access Mode:                           scsi2
  Hosts (enabled):                       phys-schost-2, phys-schost-3
```

▼ How to List the Quorum Configuration

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

You do not need to be superuser to list the quorum configuration. You can assume any role that provides `solaris.cluster.read` RBAC authorization.

Note – When you increase or decrease the number of node attachments to a quorum device, the quorum vote count is not automatically recalculated. You can reestablish the correct quorum vote if you remove all quorum devices and then add them back into the configuration. For a two-node cluster, temporarily add a new quorum device before you remove and add back the original quorum device. Then remove the temporary quorum device.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

◗ **Use** clquorum**(1CL) to list the quorum configuration.**

```
# clquorum show +
```

Example 6–9 Listing the Quorum Configuration

```
# clquorum show +

=== Cluster Nodes ===

Node Name:                            phys-schost-2
  Node ID:                            1
  Quorum Vote Count:                  1
  Reservation Key:                    0x43BAC41300000001

Node Name:                            phys-schost-3
  Node ID:                            2
  Quorum Vote Count:                  1
  Reservation Key:                    0x43BAC41300000002


=== Quorum Devices ===

Quorum Device Name:                   d3
  Enabled:                            yes
  Votes:                              1
  Global Name:                        /dev/did/rdsk/d20s2
  Type:                               scsi
```

```
Access Mode:                              scsi2
Hosts (enabled):                          phys-schost-2, phys-schost-3
```

▼ How to Repair a Quorum Device

Use this procedure to replace a malfunctioning quorum device.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Remove the disk device that you are replacing as a quorum device.

Note – If the device you intend to remove is the last quorum device, you might want to first add another disk as a new quorum device. This step assures a valid quorum device if a failure occurs during the replacement procedure. See "Adding a Quorum Device" on page 190 to add a new quorum device.

See "How to Remove a Quorum Device" on page 198 to remove a disk device as a quorum device.

2   Replace the disk device.

To replace the disk device, see the hardware procedures for the disk enclosure in the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

3   Add the replaced disk as a new quorum device.

See "Adding a Quorum Device" on page 190 to add a disk as a new quorum device.

Note – If you added an additional quorum device in Step 1, it is now safe to remove it. See "How to Remove a Quorum Device" on page 198 to remove the quorum device.

7

Administering Cluster Interconnects and Public Networks

This chapter provides the software procedures for administering the Sun Cluster interconnects and public networks.

Administering the cluster interconnects and public networks consists of both hardware and software procedures. Typically, you configure the cluster interconnects and public networks, including Internet Protocol (IP) Network Multipathing groups, when you initially install and configure the cluster. If you later need to alter a cluster interconnect network configuration, you can use the software procedures in this chapter. For information about configuring IP Network Multipathing groups in a cluster, see the section "Administering the Public Network" on page 223.

This chapter provides information and procedures for the following topics.

- "Administering the Cluster Interconnects" on page 209
- "Administering the Public Network" on page 223

For a high-level description of the related procedures in this chapter, see Table 7–1 and Table 7–3.

Refer to the *Sun Cluster Concepts Guide for Solaris OS* document for background and overview information about the cluster interconnects and public networks.

Administering the Cluster Interconnects

This section provides the procedures for reconfiguring cluster interconnects, such as cluster transport adapters and cluster transport cables. These procedures require that you install Sun Cluster software.

Most of the time, you can use the clsetup utility to administer the cluster transport for the cluster interconnects. See the clsetup(1CL) man page for more information. If you are running on the Solaris 10 OS, all cluster interconnect commands must be run in the global zone.

For cluster software installation procedures, see the *Sun Cluster Software Installation Guide for Solaris OS*. For procedures about servicing cluster hardware components, see the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

Note – You can usually choose to use the default port name, where appropriate, during cluster interconnect procedures. The default port name is the same as the internal node ID number of the node that hosts the adapter end of the cable. However, you cannot use the default port name for certain adapter types, such as SCI.

TABLE 7–1 Task List: Administering the Cluster Interconnect

Task	Instructions
Administer the cluster transport by using `clsetup`(1CL)	"How to Access the `clsetup` Utility" on page 21
Check the status of the cluster interconnect by using `clinterconnect status`	"How to Check the Status of the Cluster Interconnect" on page 211
Add a cluster transport cable, transport adapter, or switch by using `clsetup`	"How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches" on page 212
Remove a cluster transport cable, transport adapter, or transport switch by using `clsetup`	"How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches" on page 214
Enable a cluster transport cable by using `clsetup`	"How to Enable a Cluster Transport Cable" on page 217
Disable a cluster transport cable by using `clsetup`	"How to Disable a Cluster Transport Cable" on page 218
Determining an transport adapter's instance number	"How to Determine a Transport Adapter's Instance Number" on page 220
Changing the IP address or the address range of an existing cluster	"How to Change the Private Network Address or Address Range of an Existing Cluster" on page 221

Dynamic Reconfiguration With Cluster Interconnects

You must consider a few issues when completing dynamic reconfiguration (DR) operations on cluster interconnects.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support (except for the operating system quiescence operation). Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect nonnetwork IO devices during a DR detach operation.

- Sun Cluster rejects DR remove-board operations performed on active private interconnect interfaces.

- If the DR remove-board operation pertains to an active private interconnect interface, Sun Cluster rejects the operation and identifies the interface that would be affected by the operation.

⚠️ **Caution –** Sun Cluster requires that each cluster node has at least one functioning path to every other cluster node. Do not disable a private interconnect interface that supports the last path to any cluster node.

Complete the following procedures in the order indicated when performing DR operations on public network interfaces.

TABLE 7–2 Task Map: Dynamic Reconfiguration with Public Network Interfaces

Task	Instructions
1. Disable and remove the interface from the active interconnect	"Dynamic Reconfiguration With Public Network Interfaces" on page 225
2. Perform the DR operation on the public network interface.	*Sun Enterprise 10000 DR Configuration Guide*
	and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (from the *Solaris 9 on Sun Hardware*, and *Solaris 10 on Sun Hardware* collections)

▼ How to Check the Status of the Cluster Interconnect

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

You do not need to be logged in as superuser to perform this procedure.

1   Check the status of the cluster interconnect.

```
# clinterconnect status
```

2   Refer to the following table for common status messages.

Status Message	Description and Possible Action
Path online	The path is currently functioning correctly. No action is necessary.
Path waiting	The path is currently being initialized. No action is necessary.
Faulted	The path is not functioning. This can be a transient state when paths are going between the waiting and online state. If the message persists when clinterconnect status is rerun, take corrective action.

Example 7–1   Checking the Status of the Cluster Interconnect

The following example shows the status of a functioning cluster interconnect.

```
# clinterconnect status
-- Cluster Transport Paths --
                    Endpoint            Endpoint            Status
                    --------            --------            ------
  Transport path:   phys-schost-1:qfe1  phys-schost-2:qfe1  Path online
  Transport path:   phys-schost-1:qfe0  phys-schost-2:qfe0  Path online
  Transport path:   phys-schost-1:qfe1  phys-schost-3:qfe1  Path online
  Transport path:   phys-schost-1:qfe0  phys-schost-3:qfe0  Path online
  Transport path:   phys-schost-2:qfe1  phys-schost-3:qfe1  Path online
  Transport path:   phys-schost-2:qfe0  phys-schost-3:qfe0  Path online
```

▼ How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches

For information about the requirements for the cluster private transport, see "Interconnect Requirements and Restrictions" in *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Ensure that the physical cluster transport cables are installed.

For the procedure on installing a cluster transport cable, see the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

2   Become superuser on any node in the cluster.

3   Start the clsetup **utility.**

clsetup

The Main Menu is displayed.

4   Type the number that corresponds to option for displaying the Cluster Interconnect Menu.

Note – If your configuration uses SCI adapters, do not accept the default when you are prompted for the adapter connections (the port name) during the "Add" portion of this procedure. Instead, provide the port name (0, 1, 2, or 3) found on the Dolphin switch, to which the node is *physically* cabled.

5   **Type the number that corresponds to the option for adding a transport cable .**

Follow the instructions and type the requested information.

6   **Type the number that corresponds to the option for adding the transport adapter to a node.**

Follow the instructions and type the requested information.

7   **Type the number that corresponds to the option for adding the transport switch.**

Follow the instructions and type the requested information.

8   **Verify that the cluster transport cable, transport adapter, or transport switch is added.**

    # **clinterconnect show** *node:adapter,adapternode*
    # **clinterconnect show** *node:adapter*
    # **clinterconnect show** *node:switch*

Example 7–2   Adding a Cluster Transport Cable, Transport Adapter, or Transport Switch

The following example shows how to add a transport cable, transport adapter, or transport switch to a node by using the clsetup utility.

[*Ensure that the physical cable is installed.*]
[*Start the clsetup utility:*]
clsetup
[*Select Cluster interconnect*]

[*Select either Add a transport cable,*
Add a transport adapter to a node,
or Add a transport switch.}
[*Answer the questions when prompted.*]
    [*You Will Need:*   ]
[*Information: Example:*[
    node names   phys-schost-1
    adapter names   qfe2
    switch names   hub2
    transport type   dlpi
[*Verify that the* clinterconnect
 command completed successfully:]Command completed successfully.
Quit the clsetup Cluster Interconnect Menu and Main Menu.
[*Verify that the cable, adapter, and switch are added:*]
clinterconnect show phys-schost-1:qfe2,hub2

```
    ===Transport Cables ===
Transport Cable:                    phys-schost-1:qfe2@0,hub2
  Endpoint1:                  phys-schost-2:qfe0@0
  Endpoint2:                  ethernet-1@2 ???? Should this be hub2?
  State:                      Enabled
```

```
# clinterconnect show phys-schost-1:qfe2
=== Transport Adepters for qfe2
Transport Adapter:                             qfe2
  Adapter State:                               Enabled
  Adapter Transport Type:                      dlpi
  Adapter Property (device_name):              ce
  Adapter Property (device_instance):          0
  Adapter Property (lazy_free):                1
  Adapter Property (dlpi_heartbeat_timeout):   10000
  Adpater Property (dlpi_heartbeat_quantum):   1000
  Adapter Property (nw_bandwidth):             80
  Adapter Property (bandwidth):                70
  Adapter Property (ip_address):               172.16.0.129
  Adapter Property (netmask):                  255.255.255.128
  Adapter Port Names:                          0
  Adapter Port STate (0):                      Enabled
```

```
# clinterconnect show phys-schost-1:hub2
```

```
=== Transport Switches ===
Transport Switch:                              hub2
  Switch State:                                Enabled
  Switch Type:                                 switch
  Switch Port Names:                           1 2
  Switch Port State(1):                        Enabled
  Switch Port State(2):                        Enabled
```

Next Steps   To check the interconnect status of your cluster transport cable see "How to Check the Status of the Cluster Interconnect" on page 211.

▼ How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

Use the following procedure to remove cluster transport cables, transport adapters, and transport switches from a node configuration. When a cable is disabled, the two endpoints of the cable remain configured. An adapter cannot be removed if it is still in use as an endpoint on a transport cable.

⚠ **Caution –** Each cluster node needs at least one functioning transport path to every other node in the cluster. No two nodes should be isolated from one another. Always verify the status of a node's cluster interconnect before disabling a cable. Only disable a cable connection after you have verified that it is redundant. That is, ensure that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser on any node in the cluster.

2 Check the status of the remaining cluster transport path.

```
# clinterconnect status
```

⚠ **Caution –** If you receive an error such as "path faulted" while attempting to remove one node of a two-node cluster, investigate the problem before continuing with this procedure. Such a problem could indicate that a node path is unavailable. Removing the remaining operational path takes the node out of cluster membership and could result in a cluster reconfiguration.

3 Start the clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

4 Type the number that corresponds to the option for accessing the Cluster Interconnect menu.

5 Type the number that corresponds to the option for disabling the transport cable.

Follow the instructions and type the requested information. You 'need to know the applicable node names, adapter names, and switch names.

6 Type the number that corresponds to the option for removing the transport cable.

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.

Note – If you are removing a physical cable, disconnect the cable between the port and the destination device.

7 Type the number that corresponds to the option for removing the transport adapter from a node.

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.

Note – If you are removing a physical adapter from a node, see the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS* for hardware service procedures.

---

8 Type the number the corresponds to the option for removing a transport switch.

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.

Note – A switch cannot be removed if any of the ports are still in use as endpoints on any transport cables.

---

9 Verify that the cable, adapter, or switch has been removed.

```
# clinterconnect show node:adapter,adapternode
# clinterconnect show  node:adapter
# clinterconnect show  node:switch
```

The transport cable or adapter removed from the respective node should not appear in the output from this command.

Example 7–3 Removing a Transport Cable, Transport Adapter, or Transport Switch

The following example shows how to remove a transport cable, transport adapter, or transport switch by using the clsetup command.

[*Become superuser on any node in the cluster.*]
[*Start the utility:*]
 # clsetup
[*Select Cluster interconnect.*[
[*Select either Remove a transport cable,*
Remove a transport adapter to a node,
or Remove a transport switch.[
[*Answer the questions when prompted.*[
    You Will Need:
Information   Example:
   node names   phys-schost-1
   adapter names   qfe1
   switch names   hub1
[*Verify that the* clinterconnect
 command was completed successfully:]
Command completed successfully.
[*Quit the clsetup utility Cluster Interconnect Menu and Main Menu.*]
[*Verify that the cable, adapter, or switch is removed:*]
 # clinterconnect show phys-schost-1:qfe2,hub2
 ===Transport Cables ===
Transport Cable: phys-schost-2:qfe2@0,hub2
  Cable Endpoint1:   phys-schost-2:qfe0@0

```
    Cable Endpoint2:                  ethernet-1@2 ??? Should this be hub2???
    Cable State:                      Enabled

# clinterconnect show phys-schost-1:qfe2
=== Transport Adepters for qfe2
Transport Adapter:                          qfe2
  Adapter State:                              Enabled
  Adapter Transport Type:                     dlpi
  Adapter Property (device_name):             ce
  Adapter Property (device_instance):         0
  Adapter Property (lazy_free):               1
  Adapter Property (dlpi_heartbeat_timeout):  10000
  Adpater Property (dlpi_heartbeat_quantum):  1000
  Adapter Property (nw_bandwidth):            80
  Adapter Property (bandwidth):               70
  Adapter Property (ip_address):              172.16.0.129
  Adapter Property (netmask):                 255.255.255.128
  Adapter Port Names:                         0
  Adapter Port STate (0):                     Enabled

# clinterconnect show phys-schost-1:hub2
=== Transport Switches ===
Transport Switch:                           hub2
  Switch State:                               Enabled
  Switch Type:                                switch
  Switch Port Names:                          1 2
  Switch Port State(1):                       Enabled
  Switch Port State(2):                       Enabled
```

▼ How to Enable a Cluster Transport Cable

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

This option is used to enable an already existing cluster transport cable.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser on any node in the cluster.**

2 **Start the clsetup utility.**

   # **clsetup**

   The Main Menu is displayed.

3 **Type the number that corresponds to the option for accessing the Cluster Interconnect menu and press the Return key.**

4 **Type the number that corresponds to the option for enabling the transport cable and press the Return key.**

Follow the instructions when prompted. You need to provide both the node and the adapter names of one of the endpoints of the cable that you are trying to identify.

5 **Verify that the cable is enabled.**

```
# clinterconnect show  node:adapter,adapternode
```

Example 7–4 Enabling a Cluster Transport Cable

This example shows how to enable a cluster transport cable on adapter qfe-1, located on the node phys-schost-2.

[*Become superuser on any node.*]
[*Start the clsetup utility:*]
```
# clsetup
```
[*Select Cluster interconnect>Enable a transport cable.*[

[*Answer the questions when prompted.*[
[*You will need the following information.*[
    You Will Need:
Information: *Example:*
   node names   phys-schost-2
   adapter names   qfe1
   switch names   hub1
[*Verify that the* scinterconnect
 command was completed successfully:]

```
clinterconnect  enable phys-schost-2:qfe1

Command completed successfully.
```
[*Quit the clsetup Cluster Interconnect Menu and Main Menu.*]
[*Verify that the cable is enabled:*]
```
# clinterconnect show  phys-schost-1:qfe2,hub2
  Transport cable:   phys-schost-2:qfe1@0 ethernet-1@2    Enabled
  Transport cable:   phys-schost-3:qfe0@1 ethernet-1@3    Enabled
  Transport cable:   phys-schost-1:qfe0@0 ethernet-1@1    Enabled
```

▼ How to Disable a Cluster Transport Cable

You can also accomplish this procedure by using the Sun Cluster Manager GUI. See the Sun Cluster Manager online help for more information.

You might need to disable a cluster transport cable to temporarily shut down a cluster interconnect path. A temporary shutdown is useful when troubleshooting a cluster interconnect problem or when replacing cluster interconnect hardware.

When a cable is disabled, the two endpoints of the cable remain configured. An adapter cannot be removed if it is still in use as an endpoint in a transport cable.

> **Caution** – Each cluster node needs at least one functioning transport path to every other node in the cluster. No two nodes should be isolated from one another. Always verify the status of a node's cluster interconnect before disabling a cable. Only disable a cable connection after you have verified that it is redundant. That is, ensure that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser on any node in the cluster.**

2   **Check the status of the cluster interconnect before disabling a cable.**

```
# clinterconnect status
```

> **Caution** – If you receive an error such as "path faulted" while attempting to remove one node of a two-node cluster, investigate the problem before continuing with this procedure. Such a problem could indicate that a node path is unavailable. Removing the remaining operational path takes the node out of cluster membership and could result in a cluster reconfiguration.

3   **Start the** clsetup **utility.**

```
# clsetup
```

The Main Menu is displayed.

4   **Type the number that corresponds to the option for accessing the Cluster Interconnect Menu and press the Return key.**

5   **Type the number that corresponds to the option for disabling the transport cable and press the Return key.**

Follow the instructions and provide the requested information. All of the components on this cluster interconnect will be disabled. You need to provide both the node and the adapter names of one of the endpoints of the cable that you are trying to identify.

6   **Verify that the cable is disabled.**

```
# clinterconnect show  node:adapter,adapternode
```

Example 7–5   Disabling a Cluster Transport Cable

This example shows how to disable a cluster transport cable on adapter qfe-1, located on the node phys-schost-2.

[*Become superuser on any node.*]
[*Start the clsetup utility:*]
clsetup
[*Select Cluster interconnect>Disable a transport cable.*]

[*Answer the questions when prompted.*]
[*You will need the following information.*]
  [*You Will Need:*]
Information: *Example:*
    node names   phys-schost-2
    adapter names   qfe1
    switch names   hub1
[*Verify that the* clinterconnect
 command was completed successfully:]
 Command completed successfully.
[*Quit the scsetup Cluster Interconnect Menu and Main Menu.*]
[*Verify that the cable is disabled:*]
clinterconnect show -p phys-schost-1:qfe2,hub2
  Transport cable: phys-schost-2:qfe1@0 ethernet-1@2   Disabled
  Transport cable: phys-schost-3:qfe0@1 ethernet-1@3   Enabled
  Transport cable: phys-schost-1:qfe0@0 ethernet-1@1   Enabled

▼ How to Determine a Transport Adapter's Instance Number

You need to determine a transport adapter's instance number to ensure that you add and remove the correct transport adapter through the clsetup command. The adapter name is a combination of the type of the adapter and the adapter's instance number. This procedure uses an SCI-PCI adapter as an example.

1   Based on the slot number, find the adapter's name.

The following screen is an example and might not reflect your hardware.

prtdiag
```
...
========================= IO Cards =========================
                        Bus  Max
 IO  Port Bus      Freq Bus  Dev,
Type  ID  Side Slot MHz  Freq Func State Name Model
---- ---- ---- ---- ---- ---- ---- ----- --------------------------------
PCI   8    B    2    33   33  2,0   ok    pci11c8,0-pci11c8,d665.11c8.0.0
```

```
PCI   8   B   3   33  33 3,0  ok    pci11c8,0-pci11c8,d665.11c8.0.0
...
```

2. **Using the adapter's path, find the adapter's instance number.**

The following screen is an example and might not reflect your hardware.

```
# grep sci /etc/path_to_inst
"/pci@1f,400/pci11c8,o@2" 0 "sci"
"/pci@1f,4000.pci11c8,0@4 "sci"
```

3. **Using the adapter's name and slot number, find the adapter's instance number.**

The following screen is an example and might not reflect your hardware.

```
# prtconf
...
 pci, instance #0
              pci11c8,0, instance #0
            pci11c8,0, instance #1
...
```

▼ How to Change the Private Network Address or Address Range of an Existing Cluster

Use this procedure to change a private network address or the range of network addresses used or both.

1. **Reboot all cluster nodes into noncluster mode by performing the following substeps on each cluster node:**

   a. **Become superuser or assume a role that provides** solaris.cluster.admin **RBAC authorization on the cluster node to be started in noncluster mode.**

   b. **Shut down the node by using the** clnode evacuate **and** cluster shutdown **commands.**

   The clnode evacuate command switches over all device groups from the specified node to the next preferred node. The command also switches all resource groups from global or non-global zones on the specified node to the next-preferred global or non-global zones on other nodes.

   ```
   # clnode evacuate node
   # cluster shutdown -T0 -y
   ```

2. **From one node, start the** clsetup **utility.**

When run in noncluster mode, the clsetup utility displays the Main Menu for noncluster-mode operations.

3 Type the number that corresponds to the option for Change IP Address Range and press the Return key.

The `clsetup` utility displays the current private-network configuration, then asks if you would like to change this configuration.

4 To change either the private-network IP address or the IP address range, type yes **and press the Return key.**

The `clsetup` utility displays the default private-network IP address, 172.16.0.0, and asks if it is okay to accept this default.

5 Change or accept the private-network IP address.

- **To accept the default private-network IP address and proceed to changing the IP address range, type** yes **and press the Return key.**

  The `clsetup` utility will ask if it is okay to accept the default netmask. Skip to the next step to enter your response.

- **To change the default private-network IP address, perform the following substeps.**

  a. Type no **in response to the** `clsetup` **utility question about whether it is okay to accept the default address, then press the Return key.**

  The `clsetup` utility will prompt for the new private-network IP address.

  b. Type the new IP address and press the Return key.

  The `clsetup` utility displays the default netmask and then asks if it is okay to accept the default netmask.

6 Change or accept the default private-network IP address range.

The default netmask is 255.255.248.0. This default IP address range supports up to 64 nodes and up to 10 private networks in the cluster.

- **To accept the default IP address range, type** yes **and press the Return key.**
  Then skip to the next step.

- **To change the IP address range, perform the following substeps.**

  a. Type no **in response to the** `clsetup` **utility's question about whether it is okay to accept the default address range, then press the Return key.**

  When you decline the default netmask, the `clsetup` utility prompts you for the number of nodes and private networks that you expect to configure in the cluster.

  b. Enter the number of nodes and private networks that you expect to configure in the cluster.

  From these numbers, the `clsetup` utility calculates two proposed netmasks:

- The first netmask is the minimum netmask to support the number of nodes and private networks that you specified.

- The second netmask supports twice the number of nodes and private networks that you specified, to accommodate possible future growth.

    c. Specify either of the calculated netmasks, or specify a different netmask that supports the expected number of nodes and private networks.

7   **Type** yes **in response to the** clsetup **utility's question about proceeding with the update.**

8   **When finished, exit the** clsetup **utility.**

9   **Reboot each cluster node back into cluster mode by completing the following substeps for each cluster node:**

    a. Boot the node.

- On SPARC based systems, do the following:

```
ok boot
```

- On x86 based systems, do the following:

When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+----------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                               |
| Solaris failsafe                                                     |
|                                                                      |
+----------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

10   **Verify that the node has booted without error, and is online.**

```
# cluster status -t node
```

Administering the Public Network

Sun Cluster 3.1 4/04, Sun Cluster 3.1 8/05, and Sun Cluster 3.2 support the Solaris software implementation of Internet Protocol (IP) Network Multipathing for public networks. Basic IP Network Multipathing administration is the same for both cluster and noncluster environments. Multipathing administration is covered in the appropriate Solaris OS documentation. However, review the guidelines that follow before administering IP Network Multipathing in a Sun Cluster environment.

How to Administer IP Network Multipathing Groups in a Cluster

Before performing IP Network Multipathing procedures on a cluster, consider the following guidelines.

- Each public network adapter must belong to a multipathing group.
- The `local-mac-address?` variable must have a value of `true` for Ethernet adapters.
- You must configure a test IP address for each adapter in the following kinds of multipathing groups:
  - All multiple-adapter multipathing groups in a cluster that runs on the Solaris 9 or Solaris 10 OS. Single-adapter multipathing groups on the Solaris 9 or Solaris 10 OS do not require test IP addresses.
- Test IP addresses for all adapters in the same multipathing group must belong to a single IP subnet.
- Test IP addresses must not be used by normal applications because they are not highly available.
- No restrictions are placed on multipathing group naming. However, when configuring a resource group, the `netiflist` naming convention is any multipathing name followed by either the nodeID number or the node name. For example, given a multipathing group named `sc_ipmp0`, the `netiflist` naming could be either `sc_ipmp0@1` or `sc_ipmp0@phys-schost-1`, where the adapter is on the node `phys-schost-1`, which has the nodeID of `1`.
- Avoid unconfiguring (unplumbing) or bringing down an adapter of an IP Network Multipathing group without first switching over the IP addresses from the adapter to be removed to an alternate adapter in the group, using the `if_mpadm`(1M) command.
- Avoid rewiring adapters to different subnets without first removing them from their respective multipathing groups.
- Logical adapter operations can be done on an adapter even if monitoring is on for the multipathing group.
- You must maintain at least one public network connection for each node in the cluster. The cluster is inaccessible without a public network connection.
- To view the status of IP Network Multipathing groups on a cluster, use the `command.clinterconnect status` command

For more information about IP Network Multipathing, see the appropriate documentation in the Solaris OS system administration documentation set.

TABLE 7–3 Task Map: Administering the Public Network

Solaris Operating System Release	Instructions
SPARC: Solaris 9 Operating System	"IP Network Multipathing Topics" in *System Administration Guide: IP Services*

TABLE 7–3 Task Map: Administering the Public Network   *(Continued)*

Solaris Operating System Release	Instructions
Solaris 10 Operating System	"IP Network Multipathing Topics" in *System Administration Guide: IP Services*

For cluster software installation procedures, see the *Sun Cluster Software Installation Guide for Solaris OS*. For procedures about servicing public networking hardware components, see the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

Dynamic Reconfiguration With Public Network Interfaces

You must consider a few issues when completing dynamic reconfiguration (DR) operations on public network interfaces in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support (except for the operating system quiescence operation). Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect nonnetwork IO devices during a DR detach operation.

- DR remove-board operations can succeed only when public network interfaces are not active. Before removing an active public network interface, switch the IP addresses from the adapter to be removed to another adapter in the multipathing group, using the if_mpadm(1M) command.

- If you try to remove a public network interface card without having properly disabled it as an active network interface, Sun Cluster rejects the operation and identifies the interface that would be affected by the operation.

> ⚠ **Caution** – For multipathing groups with two adapters, if the remaining network adapter fails while you are performing the DR remove operation on the disabled network adapter, availability is impacted. The remaining adapter has no place to fail over for the duration of the DR operation.

Complete the following procedures in the order indicated when performing DR operations on public network interfaces.

TABLE 7–4 Task Map: Dynamic Reconfiguration With Public Network Interfaces

Task	Instructions
1. Switch the IP addresses from the adapter to be removed to another adapter in the multipathing group, using the if_mpadm	if_mpadm(1M) man page.
	The appropriate Solaris OS documentation:
	Solaris 9: "IP Network Multipathing Topics" in *System Administration Guide: IP Services*
	Solaris 10:Part VI, "IPMP," in *System Administration Guide: IP Services* in *System Administration Guide: IP Services*
2. Remove the adapter from the multipathing group by using the ifconfig command	The appropriate Solaris documentation:
	Solaris 9: "IP Network Multipathing Topics" in *System Administration Guide: IP Services*
	ifconfig(1M) man page
	Solaris 10: Part VI, "IPMP," in *System Administration Guide: IP Services* in *System Administration Guide: IP Services*
3. Perform the DR operation on the public network interface	*Sun Enterprise 10000 DR Configuration Guide* and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (from the *Solaris 9 on Sun Hardware*, and *Solaris 10 on Sun Hardware* collections)

8

Administering the Cluster

This chapter provides the procedures for administering items that affect the entire cluster.

This is a list of the sections in this chapter.

Overview of Administering the Cluster

This section describes how to perform administrative tasks for the entire cluster. The following table lists all these administrative tasks and the associated procedures. For Solaris 10 OS, unless other indicated perform cluster administrative tasks in the global zone.

TABLE 8–1 Task List: Administering the Cluster

Task	Instructions
Change the name of the cluster	"How to Change the Cluster Name" on page 228
List node IDs and their corresponding node names	"How to Map Node ID to Node Name" on page 229
Permit or deny new nodes to add themselves to the cluster	"How to Work With New Cluster Node Authentication" on page 229
Change the time for a cluster by using the Network Time Protocol (NTP)	"How to Reset the Time of Day in a Cluster" on page 231

TABLE 8–1 Task List: Administering the Cluster   *(Continued)*

Task	Instructions
Shut down a node to the OpenBoot PROM ok prompt on a SPARC based system or to the `Press any key to continue` message in a GRUB menu on an x86 based system	"SPARC: How to Display the OpenBoot PROM (OBP) on a Node" on page 233
Change the private host name	"How to Change the Node Private Host Name" on page 234
Put a cluster node in maintenance state	"How to Put a Node Into Maintenance State" on page 239
Bring a cluster node out of maintenance state	"How to Remove Node From Maintenance State" on page 240
Add a node to a cluster	"Adding a Cluster Node" on page 243
Remove a node from a cluster	"Removing a Cluster Node" on page 247
Uninstall Sun Cluster software from a node	"How to Uninstall Sun Cluster Software From a Cluster Node" on page 253
Correct error messages	"How to Correct Error Messages" on page 255

▼ How to Change the Cluster Name

If necessary, you can change the cluster name after initial installation.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser on any node in the cluster.

2   Start the `clsetup` **utility.**

    # **clsetup**

The Main Menu is displayed.

3   To change the cluster name, type the number that corresponds to the option for Other Cluster Properties.

The Other Cluster Properties menu is displayed.

4   Make your selection from the menu and follow the onscreen instructions.

Example 8–1 Changing the Cluster Name

The following example shows the cluster(1CL) command generated from the clsetup(1CL) utility to change to the new cluster name, dromedary.

```
# cluster -c dromedary
```

▼ How to Map Node ID to Node Name

During Sun Cluster installation, each node is automatically assigned a unique node ID number. The node ID number is assigned to a node in the order in which it joins the cluster for the first time. After the node ID number is assigned, the number cannot be changed. The node ID number is often used in error messages to identify which cluster node the message concerns. Use this procedure to determine the mapping between node IDs and node names.

You do not need to be superuser to list configuration information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

◗ **Use the clnode(1CL) command to list the cluster configuration information.**

```
% clnode show | grep Node
```

Example 8–2 Mapping the Node ID to the Node Name

The following example shows the node ID assignments.

```
% clnode show | grep Node
=== Cluster Nodes ===
Node Name:              phys-schost1
  Node ID:             1
Node Name:               phys-schost2
  Node ID:             2
Node Name:              phys-schost3
  Node ID:             3
```

▼ How to Work With New Cluster Node Authentication

Sun Cluster enables you to determine if new nodes can add themselves to the cluster and with what type of authentication. You can permit any new node to join the cluster over the public network, deny new nodes from joining the cluster, or indicate a specific node that can join the cluster. New nodes can be authenticated by using either standard UNIX or Diffie-Hellman (DES) authentication. If you select DES authentication, you must also configure all necessary encryption keys before a node can join. See the keyserv(1M) and publickey(4) man pages for more information.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser on any node in the cluster.

2   Start the clsetup**(1CL) utility.**

```
# clsetup
```

The Main Menu is displayed.

3   To work with cluster authentication, type the number that corresponds to the option for new nodes.

The New Nodes menu is displayed.

4   Make your selection from the menu and follow the onscreen instructions.

Example 8–3   Preventing a New Machine From Being Added to the Cluster

The following example shows the claccess command generated from the clsetup utility that would prevent new machines from being added to the cluster.

```
# claccess deny -h hostname
```

Example 8–4   Permitting All New Machines to Be Added to the Cluster

The following example shows the claccess command generated from the clsetup utility that would enable all new machines to be added to the cluster.

```
# claccess allow-all
```

Example 8–5   Specifying a New Machine to Be Added to the Cluster

The following example shows the claccess command generated from the clsetup utility to enable a single new machine to be added to the cluster.

```
# claccess allow -h hostname
```

Example 8–6   Setting the Authentication to Standard UNIX

The following example shows the claccess command generated from the clsetup utility to reset to standard UNIX authentication for new nodes that are joining the cluster.

```
# claccess set -p protocol=sys
```

Example 8–7 Setting the Authentication to DES

The following example shows the `claccess` command generated from the `clsetup` utility to use DES authentication for new nodes that are joining the cluster.

```
# claccess set -p protocol=des
```

When using DES authentication, you must also configure all necessary encryption keys before a node can join the cluster. See the `keyserv`(1M) and `publickey`(4) man pages for more information.

▼ How to Reset the Time of Day in a Cluster

Sun Cluster uses the Network Time Protocol (NTP) to maintain time synchronization between cluster nodes. Adjustments in the cluster occur automatically as needed when nodes synchronize their time. See the *Sun Cluster Concepts Guide for Solaris OS* and the *Network Time Protocol User's Guide* for more information.

⚠ **Caution** – When using NTP, do not attempt to adjust the cluster time while the cluster is up and running. Do not adjust the time by using the `date`(1), `rdate`(1M), `xntpd`(1M), or `svcadm`(1M) commands interactively or within `cron`(1M) scripts.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser on any node in the cluster.

2 Shut down the cluster.

```
# cluster shutdown -g0 -y -i 0
```

3 Verify that the node is showing the `ok` prompt on a SPARC-based system or the `Press any key to continue` message on GRUB menu on an x86 based system.

4 Boot the node in noncluster mode.

- On SPARC based systems, perform the following command:

  ```
  ok boot -x
  ```

- On x86 based system, perform the following commands:

```
phys-schost# shutdown -g -y -i0

Press any key to continue
```

a. In the GRUB menu, use the arrow keys to select the appropriate Solaris entry and type e **to edit its commands.**

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-------------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                                   |
| Solaris failsafe                                                         |
|                                                                         |
+-------------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration* in *System Administration Guide: Basic Administration*.

b. In the boot parameters screen, use the arrow keys to select the kernel entry and type e **to edit the entry.**

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-------------------------------------------------------------------------+
| root (hd0,0,a)                                                           |
| kernel /platform/i86pc/multiboot                                        |
| module /platform/i86pc/boot_archive                                     |
+-------------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

c. Add -x **to the command to specify that the system boot into noncluster mode.**

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. ]

grub edit> kernel /platform/i86pc/multiboot -x
```

d. Press the Enter key to accept the change and return to the boot parameters screen.

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+----------------------------------------------------------------------+
| root (hd0,0,a)                                                       |
| kernel /platform/i86pc/multiboot -x                                 |
| module /platform/i86pc/boot_archive                                 |
+----------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.-
```

e. Type b **to boot the node into noncluster mode.**

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -x option to the kernel boot parameter command.

---

5   On a single node, set the time of day by running the date **command.**

date *HHMM.SS*

6   On the other machines, synchronize the time to that node by running the rdate(1M) **command.**

rdate *hostname*

7   Boot each node to restart the cluster.

reboot

8   Verify that the change occurred on all cluster nodes.

On each node, run the date command.

date

▼ SPARC: How to Display the OpenBoot PROM (OBP) on a Node

Use this procedure if you need to configure or change OpenBoot PROM settings.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Connect to the console on the node to be shutdown.

telnet *tc_name* *tc_port_number*

tc_name   Specifies the name of the terminal concentrator.

tc_*port_number*   Specifies the port number on the terminal concentrator. Port numbers are
                               configuration dependent. Typically, ports 2 and 3 (5002 and 5003) are used for
                               the first cluster installed at a site.

2   Shut down the cluster node gracefully by using the clnode evacuate **command then the** shutdown
command. The clnode evacuate **command switches over all device groups from the specified node
to the next preferred node. The command also switches all resource groups from global or
non-global zones on the specified node to the next-preferred global or non-global zones on other
nodes.**

clnode evacuate *node*
shutdown -g0 -y

⚠️   **Caution –** Do not use send brk on a cluster console to shut down a cluster node.

---

3   Execute the OBP commands.

▼ How to Change the Node Private Host Name

Use this procedure to change the private host name of a cluster node after installation has been
completed.

Default private host names are assigned during initial cluster installation. The default private host
name takes the form clusternode< *nodeid*>-priv, for example: clusternode3-priv . Change a
private host name only if the name is already in use in the domain.

⚠️   **Caution –** Do not attempt to assign IP addresses to new private host names. The clustering software
assigns them.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have
short forms. Except for the long and short forms of the command names, the commands are
identical. For a list of the commands and their short forms, see Appendix A.

**1   Disable, on all nodes in the cluster, any data service resources or other applications that might cache
private host names.**

clresource disable resource[,...]

Include the following in the applications you disable.

- HA-DNS and HA-NFS services, if configured
- Any application that has been custom configured to use the private host name
- Any application that is being used by clients over the private interconnect

See the clresource(1CL) man page and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for information about using the clresource command.

2   If your NTP configuration file refers to the private host name that you are changing, bring down the Network Time Protocol (NTP) daemon on each node of the cluster.

- SPARC: If you are using Solaris 9 OS, use the xntpd command to shut down the Network Time Protocol (NTP) daemon. See the xntpd(1M) man page for more information about the NTP daemon.

  # **/etc/init.d/xntpd.cluster stop**

- If you are using Solaris 10 OS, use the svcadm command to shut down the Network Time Protocol (NTP) daemon. See the svcadm(1M) man page for more information about the NTP daemon.

  # **svcadm disable ntp**

3   Run the clsetup(1CL) utility to change the private host name of the appropriate node.

To need to run the utility only from one of the nodes in the cluster.

Note – When selecting a new private host name, ensure that the name is unique to the cluster node.

---

4   Type the number that corresponds to the option for private host name.

5   Type the number that corresponds to the option for changing a private host name.

Answer the questions when prompted. You are asked the name of the node whose private host name is being changed (clusternode< *nodeid*>-priv), and the new private host name.

6   Flush the name service cache.

Perform this step on each node in the cluster. Flushing prevents the cluster applications and data services from trying to access the old private host name.

nscd -i hosts

7   If you changed a private host name in your NTP configuration file, update your NTP configuration file (ntp.conf **or** ntp.conf.cluster**) on each node.**

a.   Use the editing tool of your choice.

If you perform this step at installation, also remember to remove names for nodes which are configured. The default template is preconfigured with 16 nodes. Typically, the ntp.conf.cluster file is identical on each cluster node.

   b. Verify that you can successfully ping the new private host name from all cluster nodes.

   c. Restart the NTP daemon.

   Perform this step on each node of the cluster.

   - SPARC: If you are using Solaris 9 OS, use the `xntpd` command to restart the NTP daemon.

     If you are using the `ntp.conf.cluster` file, type the following:

     `# /etc/init.d/xntpd.cluster start`

     If you are using the `ntp.conf` file, type the following:

     `# /etc/init.d/xntpd start`

   - If you are using Solaris 10 OS, use the `svcadm` command to restart the NTP daemon.

     `# svcadm enable ntp`

8 Enable all data service resources and other applications that were disabled in Step 1.

`# clresource disable resource[,...]`

See the `clresource`(1CL) man page and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for information about using the `scswitch` command.

Example 8–8   Changing the Private Host Name

The following example changes the private host name from `clusternode2-priv` to `clusternode4-priv`, on node `phys-schost-2`.

```
[Disable all applications and data services as necessary.]
phys-schost-1# /etc/init.d/xntpd stop
phys-schost-1# clnode show | grep node
 ...
 private hostname:                          clusternode1-priv
 private hostname:                          clusternode2-priv
 private hostname:                          clusternode3-priv
 ...
phys-schost-1# clsetup
phys-schost-1# nscd -i hosts
phys-schost-1# vi /etc/inet/ntp.conf
 ...
 peer clusternode1-priv
 peer clusternode4-priv
 peer clusternode3-priv
phys-schost-1# ping clusternode4-priv
phys-schost-1# /etc/init.d/xntpd start
[Enable all applications and data services disabled at the beginning of the procedure.]
```

▼ How to Add a Private Host Name for a Non-Global Zone

Use this procedure to add a private host name for a cluster non-global zone after installation has been completed.

1 **Run the** `clsetup`**(1CL) utility to add a private host name on the appropriate zone.**

2 **Type the number that corresponds to the option for private host names and press the Return key.**

3 **Type the number that corresponds to the option for adding a zone private host name and press the Return key.**

Answer the questions when prompted. There is no default for a non-global zone private host name. You will need to provide a host name.

▼ How to Change the Non-Global Zone Private Host Name

Use this procedure to change the private host name of a cluster non-global zone after installation has been completed.

Private host names are assigned during initial cluster installation. The private host name takes the form clusternode< *nodeid*>-priv, for example: clusternode3-priv . Change a private host name only if the name is already in use in the domain.

⚠ **Caution –** Do not attempt to assign IP addresses to new private host names. The clustering software assigns them.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Disable, on all nodes in the cluster, any data service resources or other applications that might cache private host names.**

```
# clresource disable resource1, resource2
```

Include the following in the applications you disable.

- HA-DNS and HA-NFS services, if configured
- Any application that has been custom configured to use the private host name
- Any application that is being used by clients over the private interconnect

See the clresource(1CL) man page and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for information about using the clresource command.

2   **Run the** clsetup**(1CL) utility to change the private host name of the appropriate non-global zone.**

    You need to perform this step only from one of the nodes in the cluster.

    Note – When selecting a new private host name, ensure that the name is unique to the cluster.

    ---

3   **Type the number that corresponds to the option for private host names and press the Return key.**

4   **Type the number that corresponds to the option for adding a zone private hostname and press the Return key.**

    There is no default for a non-global zone private host name. You will need to provide a host name.

5   **Type the number that corresponds to the option for changing a zone private hostname.**

    Answer the questions when prompted. You are asked for the name of the non-global zone whose private host name is being changed (clusternode< *nodeid*>-priv), and the new private host name.

6   **Flush the name service cache.**

    Perform this step on each node in the cluster. Flushing prevents the cluster applications and data services from trying to access the old private host name.

    ```
    # nscd -i hosts
    ```

7   **Enable all data service resources and other applications that were disabled in Step 1.**

▼ How to Delete the Private Host Name for a Non-Global Zone

Use this procedure to delete a private host name for a cluster non-global zone.

1   **Run the** clsetup**(1CL) utility to delete a private host name on the appropriate zone.**

2   **Type the number that corresponds with the option for zone private host name.**

3   **Type the number that corresponds with the option for deleting a zone private host names.**

4   **Type the name of the non-global zone private host name you are deleting.**

▼ How to Put a Node Into Maintenance State

Put a cluster node into maintenance state when taking the node out of service for an extended period of time. This way, the node does not contribute to the quorum count while it is being serviced. To put a cluster node into maintenance state, the node must be shut down by using clnode(1CL) evacuate and cluster(1CL) shutdown commands.

Note – Use the Solaris shutdown command to shut down a single node. Use the cluster shutdown command only when shutting down an entire cluster.

When a cluster node is shut down and put in maintenance state, all quorum devices that are configured with ports to the node have their quorum vote counts decremented by one. The node and quorum device vote counts are incremented by one when the node is removed from maintenance mode and brought back online.

Use the clquorum(1CL) disable command to put a cluster node to maintenance state.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on the node that you are putting into maintenance state.**

2   **Evacuate any resource groups and device groups from the node. The** clnode evacuate **command switches over all resource groups and device groups, including all non-global zones from the specified node to the next preferred node.**

    # **clnode evacuate** *node*

3   **Shut down the node that you evacuated.**

    # **shutdown -g0 -y-i 0**

4   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization on another node in the cluster and put the node that you shut down in Step 3 in maintenance state.**

    # **clquorum disable** *type* [*node*]

    type   Specifies the quorum device type. The type can be node, scsi, netapp_nas, or
                       quorum_server.

    node   Specifies the name of a node you want to put into maintenance mode.

5   **Verify that the cluster node is now in maintenance state.**

    # **clquorum status -t** *node*

The node that you put in maintenance state should have a Status of offline and 0 (zero) for Present and Possible quorum votes.

Example 8–9 Putting a Cluster Node In Maintenance State

The following example puts a cluster node in maintenance state and verifies the results. The clnode status output shows the Node votes for phys-schost-1 to be 0 (zero) and the status to be Offline. The Quorum Summary should also show reduced vote counts. Depending on your configuration, the Quorum Votes by Device output might indicate that some quorum disk devices are offline as well.

```
[On the node to be put into maintenance state:]
phys-schost-1# clnode evacute phys-schost-1
phys-schost-1# shutdown -g0 -y -i0

[On another node in the cluster:]
phys-schost-2# clquorum disable -t node phys-schost-1
phys-schost-2# clquorum status -t node

-- Quorum Votes by Node --

Node Name          Present      Possible       Status
---------          -------      --------       ------
phys-schost-1      0            0              Offline
phys-schost-2      1            1              Online
phys-schost-3      1            1              Online
```

See Also To bring a node back online, see "How to Remove Node From Maintenance State" on page 240.

▼ How to Remove Node From Maintenance State

Use the following procedure to bring a node back online and reset the quorum vote count to the default. For cluster nodes, the default quorum count is one. For quorum devices, the default quorum count is *N*-1, where *N* is the number of nodes with nonzero vote counts that have ports to the quorum device.

When a node has been put in maintenance state, the node's quorum vote count is decremented by one. All quorum devices that are configured with ports to the node will also have their quorum vote counts decremented. When the quorum vote count is reset and a node removed from maintenance state, both the node's quorum vote count and the quorum device vote count are incremented by one.

Run this procedure any time a node has been put in maintenance state and you are removing it from maintenance state.

> ⚠ **Caution –** If you do not specify either the `globaldev` or `node` options, the quorum count is reset for the entire cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1. **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization on any node of the cluster other than the one in maintenance state.**

2. **Depending on the number of nodes you have in your cluster configuration, perform one of the following steps:**
   - If you have two nodes in your cluster configuration, go to Step 4.
   - If you have more than two nodes in your cluster configuration, go to Step 3.

3. **If the node you are removing from maintenance state will have quorum devices, reset the cluster quorum count from a node other than the one in maintenance state.**

   You must reset the quorum count from a node other than the node in maintenance state before rebooting the node, or the node might hang waiting for quorum.

   ```
   # clquorum reset
   ```

   reset   The change flag that resets quorum.

4. **Boot the node that you are removing from maintenance state.**

5. **Verify the quorum vote count.**

   ```
   # clquorum status
   ```

   The node that you removed from maintenance state should have a status of `online` and show the appropriate vote count for `Present` and `Possible` quorum votes.

Example 8–10   Removing a Cluster Node From Maintenance State and Resetting the Quorum Vote Count

The following example resets the quorum count for a cluster node and its quorum devices to their defaults and verifies the result. The `scstat -q` output shows the `Node votes` for phys-schost-1 to be 1 and the status to be `online`. The `Quorum Summary` should also show an increase in vote counts.

```
phys-schost-2# clquorum reset
```

- On SPARC based systems, do the following:

  ```
  ok boot
  ```

■ On x86 based systems, do the following:

When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                                 |
| Solaris failsafe                                                       |
|                                                                       |
+-----------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

```
phys-schost-1# clquorum status

--- Quorum Votes Summary ---

          Needed   Present   Possible
          ------   -------   --------
          4        6         6


--- Quorum Votes by Node ---

Node Name       Present       Possible       Status
---------       -------       --------       ------
phys-schost-2   1             1              Online
phys-schost-3   1             1              Online


--- Quorum Votes by Device ---

Device Name         Present       Possible       Status
-----------         -------       --------       ------
/dev/did/rdsk/d3s2  1             1              Online
/dev/did/rdsk/d17s2 0             1              Online
/dev/did/rdsk/d31s2 1             1              Online
'
```

Adding a Cluster Node

This section provides instructions on adding a cluster node. The following table lists the tasks to perform when adding a node to an existing cluster. To complete the procedure correctly, these tasks must be performed in the order shown.

TABLE 8–2 Task Map: Adding a Cluster Node to an Existing Cluster

Task	Instructions
Install the host adapter on the node and verify that the existing cluster interconnects can support the new node	*Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*
Add shared storage	*Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*
Prepare the cluster for additional nodes	"How to Prepare the Cluster for Additional Cluster Nodes" in *Sun Cluster Software Installation Guide for Solaris OS* in *Sun Cluster Software Installation Guide for Solaris OS*
Add the node to the authorized node list by using `clsetup`	"How to Add a Node to the Authorized Node List" on page 243
Install and configure the software on the new cluster node	Chapter 2, "Installing Software on the Cluster," in *Sun Cluster Software Installation Guide for Solaris OS*

▼ How to Add a Node to the Authorized Node List

Before adding a machine to an existing cluster, ensure that the node has all of the necessary hardware correctly installed and configured, including an operational physical connection to the private cluster interconnect.

For hardware installation information, refer to the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS* or the hardware documentation that shipped with your server.

This procedure enables a machine to install itself into a cluster by adding its node name to the list of authorized nodes for that cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   On a current cluster member, become superuser on a current cluster member.

2   Ensure that you have correctly completed all prerequisite hardware installation and configuration tasks listed in the task map for "Adding a Cluster Node" on page 243.

3 Start the `clsetup` **utility.**

clsetup

The Main Menu is displayed.

4 Type the number that corresponds to the option for displaying the New Nodes Menu and press the Return key.

5 Type the number that corresponds to the option to modify the authorized list and press the Return key. Specify the name of a machine that can add itself.

Follow the prompts to add the node's name to the cluster. You are asked for the name of the node to be added.

6 Verify that the task has been performed successfully.

The `clsetup` utility prints a "Command completed successfully" message if it completes the task without error.

7 To prevent any new machines from being added to the cluster, type the number that corresponds to option to instruct the cluster to ignore requests to add new machines and press the Return key.

Follow the `clsetup` prompts. This option tells the cluster to ignore all requests over the public network from any new machine trying to add itself to the cluster.

8 Quit the `clsetup` **utility.**

9 Install and configure the software on the new cluster node.

Use either `scinstall` or JumpStart™ to complete the installation and configuration of the new node, as described in the *Sun Cluster Software Installation Guide for Solaris OS*.

Example 8–11 Adding a Cluster Node to the Authorized Node List

The following example shows how to add a node named phys-schost-3 to the authorized node list in an existing cluster.

[*Become superuser and execute the clsetup utility.*]
clsetup
[*Select New nodes>Specify the name of a machine which may add itself.]*
[*Answer the questions when prompted.]*
[*Verify that the scconf command completed successfully.]*

claccess allow -h phys-schost-3

    Command completed successfully.
[*Select Prevent any new machines from being added to the cluster.]*
[*Quit the clsetup New Nodes Menu and Main Menu.*]
[*Install the cluster software.*]

See Also For an overall list of tasks for adding a cluster node, see Table 8–2, "Task Map: Adding a Cluster Node."

To add a node to an existing resource group, see the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS.*

Administering a Non-Global Zone on a Node

This section provides the following information and procedures to create a non-global zone, simply referred to as a *zone*, on a cluster node.

▼ How to Create a Non-Global Zone on a Node

1 **Become superuser on the node on which you are creating the non-global zone.**

You must be in the global zone.

2 **For the Solaris 10 OS, verify on each node that multi-user services for the Service Management Facility (SMF) are online.**

If services are not yet online for a node, wait until the state becomes online before you proceed to the next step.

```
phys-schost# svcs multi-user-server
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

3 **Configure, install, and boot the new zone.**

Note – You must set the autoboot property to true to support resource-group functionality in the non-global zone.

Follow procedures in the following documentation:

a. **Perform procedures in Chapter 18, "Planning and Configuring Non-Global Zones (Tasks)," in** *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones.*

b. **Perform procedures in "Installing and Booting Zones" in** *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones.*

c. **Perform procedures in "How to Boot a Zone" in** *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones.*

4 **Verify that the zone is in the** ready **state.**

```
phys-schost# zoneadm list -v
ID  NAME      STATUS       PATH
```

```
0  global   running    /
1  my-zone  ready      /zone-path
```

5 **(Optional) Assign a private IP address and a private hostname to the zone.**

The following command chooses and assigns an available IP address from the cluster's private IP address range. The command also assigns the specified private hostname, or host alias, to the zone and maps it to the assigned private IP address.

phys-schost# **clnode set -p zprivatehostname=**hostalias node:zone

-p	Specifies a property.
zprivatehostname=hostalias	Specifies the zone private hostname, or host alias.
node	The name of the node.
zone	The name of the non-global zone.

6 **Perform the initial internal zone configuration.**

Follow the procedures in "Performing the Initial Internal Zone Configuration" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*. Choose either of the following methods:

- Log in to the zone
- Use an /etc/sysidcfg file

7 **In the non-global zone, modify the** nsswitch.conf **file.**

You must make these changes to enable the zone to resolve searches for cluster-specific hostnames and IP addresses.

a. **Log in to the zone.**

phys-schost# **zogin -c** zonename

b. **Open the** /etc/nsswitch.conf **file for editing.**

phys-schost# **vi /etc/nsswitch.conf**

c. **Add the** cluster **switch to the beginning of the lookups for the** hosts **and** netmasks **entries.**

The modified entries would appear similar to the following:

```
...
hosts:      cluster files nis [NOTFOUND=return]
...
netmasks:   cluster files nis [NOTFOUND=return]
...
```

▼ How to Remove a Non-Global Zone on a Node

1   Become superuser on the node on which you will create the non-global zone.

2   Delete the non-global zone from the system.

Follow the procedures in "Deleting a Non-Global Zone From the System" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*

Removing a Cluster Node

This section provides instructions on how to remove a cluster node. The following table lists the tasks to perform to remove a node from an existing cluster. To complete the procedure correctly, the tasks must be performed in the order shown.

⚠️   **Caution –** Do not use this procedure if your cluster is running an OPS configuration. At this time, removing a node in an OPS configuration might cause nodes to panic at reboot.

TABLE 8–3 Task Map: Removing a Cluster Node (5/02)

Task	Instructions
Move all resource groups and device groups off of the node to be removed. - Use `clnode`(1CL) evacuate	`# clnode evacuate` *node*
Remove the node from all resource groups. - Use `clresourcegroup`(1CL) remove-node	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
Remove node from all device groups - Use `cldevicegroup`(1CL) remove node, `metaset`(1M), and `clsetup`(1CL)	"How to Remove a Node From a Device Group (Solaris Volume Manager)" on page 141 "SPARC: How to Remove a Node From a Device Group (VERITAS Volume Manager)" on page 157 "How to Remove a Node From a Raw-Disk Device Group" on page 159 **Caution: If the number of desired secondaries is configured as 2 or more, it must be decreased to 1.**

TABLE 8–3 Task Map: Removing a Cluster Node (5/02)   *(Continued)*

Task	Instructions
Remove all fully connected quorum devices. - Use `clsetup`.	**Caution: Do not remove the quorum device if you are removing a node from a two-node cluster.** "How to Remove a Quorum Device" on page 198 Note that although you must remove the quorum device before you remove the storage device in the next step, you can add the quorum device back immediately afterward.
Remove all fully connected storage devices from the node. - Use `devfsadm`(1M), `\cldevice`(1CL) `refresh`.	**Caution: Do not remove the quorum device if you are removing a node from a two-node cluster.** "How to Remove Connectivity Between an Array and a Single Node, in a Cluster With Greater Than Two-Node Connectivity" on page 250
Add back the quorum devices (to only the nodes that are intended to remain in the cluster). - Use `clsetup`	"Adding a Quorum Device" on page 190
Put the node being removed into maintenance state. - Use `clnode`(1CL) evacuate, `cluster`(1CL) shutdown, and `clquorum`(1CL) disable	"How to Put a Node Into Maintenance State" on page 239
Remove all logical transport connections (transport cables and adapters) from the node being removed. - Use `clsetup`.	"How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches" on page 214
Remove all quorum devices connected to the node being removed. - Use `cluster set`, `clquorum remove`.	"How to Remove the Last Quorum Device From a Cluster" on page 199
Remove node from the cluster software configuration. - Use `clnode remove`.	"How to Remove a Node From the Cluster Software Configuration" on page 248

▼ How to Remove a Node From the Cluster Software Configuration

Perform this procedure to remove a node from the cluster.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Ensure that you have correctly completed all prerequisite tasks listed in the "Removing a Cluster Node" task map in "Adding a Cluster Node" on page 243.

Note – Ensure that you have removed the node from all resource groups, device groups, and quorum device configurations and put it into maintenance state before you continue with this procedure.

---

2 Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on a node in the cluster** *other than the node to remove***.**

3 Remove the node from the cluster.

   # **clnode remove** *nodename*

4 Verify the node removal by using scstat**(1M).**

   # **clnode status** *nodename*

5 If you intend to uninstall the Sun Cluster software from the removed node, proceed to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 253. If you do not intend to uninstall the Sun Cluster software from the remove node, you can physically remove the node from the cluster by removing the hardware connections as described in *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS***.**

Example 8–12   Removing a Node From the Cluster Software Configuration

This example shows how to remove a node (phys-schost-2) from a cluster. All commands are run from another node of the cluster (phys-schost-1).

```
[Remove the node from the cluster:]
phys-schost-1# clnode remove phys-schost-2
[Verify node removal:]
phys-schost-1# clnode status phys-schost-2
-- Cluster Nodes --
                  Node name          Status
                  ---------          ------
  Cluster node:   phys-schost-1      Online
```

See Also   To uninstall Sun Cluster software from the removed node, see "How to Uninstall Sun Cluster Software From a Cluster Node" on page 253.

For hardware procedures, see the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

For an overall list of tasks for removing a cluster node, see Table 8–3.

To add a node to an existing cluster, see "How to Add a Node to the Authorized Node List" on page 243.

▼ How to Remove Connectivity Between an Array and a Single Node, in a Cluster With Greater Than Two-Node Connectivity

Use this procedure to detach a storage array from a single cluster node, in a cluster that has three-node or four-node connectivity.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Back up all database tables, data services, and volumes that are associated with the storage array that you are removing.

2 Determine the resource groups and device groups that are running on the node to be disconnected.

```
# clresourcegroup status
# cldevicegroup status
```

3 If necessary, move all resource groups and device groups off the node to be disconnected.

⚠ **Caution (SPARC only) –** If your cluster is running Oracle Parallel Server/Oracle RAC software, shut down the Oracle Parallel Server/Oracle RAC database instance that is running on the node before you move the groups off the node. For instructions see the *Oracle Database Administration Guide*.

```
# clnode evacuate node
```

The clnode evacuate command switches over all device groups from the specified node to the next preferred node. The command also switches all resource groups from global or non-global zones on the specified node to the next-preferred global or non-global zones on other nodes.

4 Put the device groups into maintenance state.

For the procedure on acquiescing I/O activity to Veritas shared disk groups, see your VxVM documentation.

For the procedure on putting a device group in maintenance state, see the Chapter 8.

5 Remove the node from the device groups.

- If you use VxVM or raw disk, use the cldevicegroup(1CL)command to remove the device groups.
- If you use Solstice DiskSuite, use the metaset command to remove the device groups.

6 For each resource group that contains an `HAStoragePlus` **resource, remove the node from the resource group's node list.**

`# clresourcegroup remove-node` *-z zone* `-n` *node* `+ |` *resourcegroup*

node   The name of the node.

zone   The name of the non-global zone on *node* that can master the resource group. Specify *zone* only if you specified a non-global zone when you created the resource group.

See the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for more information about changing a resource group's nodelist.

Note – Resource type, resource group, and resource property names are case sensitive when `clresourcegroup` is executed.

---

7 If the storage array you are removing is the last storage array that is connected to the node, disconnect the fiber-optic cable between the node and the hub or switch that is connected to this storage array (otherwise, skip this step).

8 If you are removing the host adapter from the node that you are disconnecting, shut down and power off the node. If you are removing the host adapter from the node that you are disconnecting, skip to Step 11.

9 Remove the host adapter from the node.

For the procedure on removing host adapters, see the documentation that shipped with your node.

10 Without booting the node to, power on the node.

11 Boot the node in noncluster mode.

- SPARC: Type:

  ok **boot -x**

- x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. On the Current Boot Parameters screen, type b or i.

```
                    <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults
```

```
                        <<< timeout in 5 seconds >>>
         Select (b)oot or (i)nterpreter: b -x
```

⚠️ **Caution (SPARC only)** – The node must be in noncluster mode before you remove Oracle Parallel Server/Oracle RAC software in the next step or the node panics and potentially causes a loss of data availability.

12 SPARC: If Oracle Parallel Server/Oracle RAC software has been installed, remove the Oracle Parallel Server/Oracle RAC software package from the node that you are disconnecting.

```
# pkgrm SUNWscucm
```

⚠️ **Caution (SPARC only)** – If you do not remove the Oracle Parallel Server/Oracle RAC software from the node that you disconnected, the node panics when the node is reintroduced to the cluster and potentially causes a loss of data availability.

13 Boot the node in cluster mode.

- On SPARC based systems, do the following:

  ```
  ok boot
  ```

- On x86 based systems, do the following:

  When the GRUB menu is displayed, select the appropriate Solaris entry and press Enter. The GRUB menu appears similar to the following:

  ```
  GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
  +-----------------------------------------------------------------------+
  | Solaris 10 /sol_10_x86                                                 |
  | Solaris failsafe                                                       |
  |                                                                       |
  +-----------------------------------------------------------------------+
  Use the ^ and v keys to select which entry is highlighted.
  Press enter to boot the selected OS, 'e' to edit the
  commands before booting, or 'c' for a command-line.
  ```

14 On the node, update the device namespace by updating the /devices and /dev entries.

```
# devfsadm -C
# cldevice refresh
```

15 Bring the device groups back online.

For procedures about bringing a VERITAS shared disk group online, see your VERITAS Volume Manager documentation.

For the procedure on bringing a device group online, see the procedure on putting a device group into maintenance state.

▼ How to Uninstall Sun Cluster Software From a Cluster Node

Perform this procedure to uninstall Sun Cluster software from a cluster node before you disconnect it from a fully established cluster configuration. You can use this procedure to uninstall software from the last remaining node of a cluster.

Note – To uninstall Sun Cluster software from a node that has not yet joined the cluster or is still in installation mode, do not perform this procedure. Instead, go to "How to Uninstall Sun Cluster Software to Correct Installation Problems" in the *Sun Cluster Software Installation Guide for Solaris OS*.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Ensure that you have correctly completed all prerequisite tasks in the task map for removing a cluster node.

See Table 8–3.

Note – Ensure that you have removed the node from the cluster configuration using clnode remove before you continue with this procedure.

---

2   Become superuser on an active cluster member *other than* **the node you are uninstalling.**

3   From the active cluster member, add the node that you intend to uninstall to the cluster's node authentication list.

claccess allow -h *hostname*

-h   Specifies the name of the node to be added to the node's authentication list.

Alternately, you can use the clsetup(1CL) utility. See "How to Add a Node to the Authorized Node List" on page 243 for procedures.

4   Become superuser on the node to uninstall.

5   Reboot the node into noncluster mode.

- SPARC: Type:

```
# shutdown -g0 -y -i0ok boot -x
```

- x86: Type:

```
# shutdown -g0 -y -i0
...
                     <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -x
```

6 **In the** /etc/vfstab **file, remove all globally mounted file system entries** *except* **the**
   /global/.devices **global mounts.**

7 **If you intend to reinstall Sun Cluster software on this node, remove the Sun Cluster entry from the**
   Sun Java Enterprise System (Java ES) product registry.

   If the Java ES product registry contains a record that Sun Cluster software was installed, the Java ES
   installer shows the Sun Cluster component grayed out and does not permit reinstallation.

   a. **Start the Java ES uninstaller.**

       Run the following command, where *ver* is the version of the Java ES distribution from which you
       installed Sun Cluster software.

       ```
       # /var/sadm/prod/SUNWentsysver/uninstall
       ```

   b. **Follow the prompts to select** Sun Cluster **to uninstall.**

       For more information about using the uninstall command, see <<olink to "Uninstalling
       Components" in Java ES Installation Guide>>.

8 **If you do not intend to reinstall Sun Cluster on this cluster, disconnect the transport cables and the**
   transport switch, if any, from the other cluster devices.

   a. **If the uninstalled node is connected to a storage device that uses a parallel SCSI interface, install a**
       SCSI terminator to the open SCSI connector of the storage device after you disconnect the
       transport cables.

       If the uninstalled node is connected to a storage device that uses Fibre Channel interfaces, no
       termination is necessary.

b. Follow the documentation that shipped with your host adapter and server for disconnection procedures.

▼ How to Correct Error Messages

To correct the error messages in the previous sections, perform this procedure.

1 Attempt to rejoin the node to the cluster.

   # **boot**

2 Did the node successfully rejoin the cluster?

   ■ If no, proceed to Step 3.
   ■ If yes, perform the following steps to remove the node from device groups.

   a. If the node successfully rejoins the cluster, remove the node from the remaining device group or groups.

   Follow procedures in "How to Remove a Node From All Device Groups" on page 140.

   b. After you remove the node from all device groups, return to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 253 and repeat the procedure.

3 If the node could not rejoin the cluster, rename the node's /etc/cluster/ccr **file to any other name you choose, for example,** ccr.old**.**

   # **mv /etc/cluster/ccr /etc/cluster/ccr.old**

4 Return to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 253 and repeat the procedure.

Troubleshooting a Node Uninstallation

This section describes error messages you might receive when you run the scinstall -r command and the corrective actions to take.

Unremoved Cluster File-System Entries

The following error messages indicate that the node you removed still has cluster file systems referenced in its vfstab file.

```
Verifying that no unexpected global mounts remain in /etc/vfstab ... failed
scinstall:  global-mount1 is still configured as a global mount.
scinstall:  global-mount1 is still configured as a global mount.
scinstall:  /global/dg1 is still configured as a global mount.
```

```
scinstall:  It is not safe to uninstall with these outstanding errors.
scinstall:  Refer to the documentation for complete uninstall instructions.
scinstall:  Uninstall failed.
```

To correct this error, return to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 253 and repeat the procedure. Ensure that you successfully complete Step 6 in the procedure before you rerun the scinstall -r command.

Unremoved Listing in Device Groups

The following error messages indicate that the node you removed is still listed with a device group.

```
Verifying that no device services still reference this node ... failed
scinstall:  This node is still configured to host device service "
service".
scinstall:  This node is still configured to host device service "
service2".
scinstall:  This node is still configured to host device service "
service3".
scinstall:  This node is still configured to host device service "
dg1".

scinstall:  It is not safe to uninstall with these outstanding errors.
scinstall:  Refer to the documentation for complete uninstall instructions.
scinstall:  Uninstall failed.
```

Creating, Setting Up, and Managing the Sun Cluster SNMP Event MIB

This section describes how to create, set up, and manage the Simple Network Management Protocol (SNMP) event Management Information Base (MIB). This section also describes how to enable, disable, and change the Sun Cluster SNMP event MIB.

The Sun Cluster software currently supports one MIB, the event MIB. The event MIB notifies SNMP manager software of cluster events in real time. When enabled, the event MIB automatically sends trap notifications to all hosts that are defined by the clsnmphost command. Because clusters generate numerous notifications, only events with a severity of warning or greater are sent as trap notifications. The MIB maintains a read-only table of the most current 50 events for which a trap has been sent. This information does not persist across reboots.

Creating, setting up, and managing a Sun Cluster SNMP event MIB can involve the following tasks:

TABLE 8–4 Task Map: Creating, Setting Up, and Managing the Sun Cluster SNMP Event MIB

Task	Instructions
Add an SNMP host to the list of hosts that will receive trap notifications for the MIBs	"How to Add an SNMP Host on the Current Node" on page 257 and "How to Add an SNMP Host on Another Node" on page 258
Remove an SNMP host	"How to Remove an SNMP Host on the Current Node" on page 258 and "How to Remove an SNMP Host on Another Node" on page 259
Change an SNMP event MIB	"How to Change an SNMP Event MIB" on page 259
Enable an SNMP event MIB	"How to Enable an SNMP Event MIB" on page 260
Disable an SNMP event MIB	"How to Disable an SNMP Event MIB" on page 260
Add an SNMP user	"How to Add an SNMP User" on page 261
Remove an SNMP user	"How to Remove an SNMP User" on page 262

▼ How to Add an SNMP Host on the Current Node

This procedure shows how to add an SNMP host on the current node to the list of hosts that will receive trap notifications for the MIBs.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization.**

2 **Add the host to the SNMP host list of a community on the current node.**

```
phys-schost-1# clsnmphost add -c SNMPcommunity host
```

-c *SNMPcommunity*
   Specifies the SNMP community name that is used in conjunction with the host name.

   You must specify the SNMP community name (with the -c option) when you add a host to a community (other than public). If the specified community name does not exist, this command creates the community.

   If you use the add subcommand without the -c option, the subcommand uses public as the default community name.

host
   Specifies the name, IP address, or IPv6 address of the host that is provided access to the SNMP MIBs in the cluster.

▼ How to Add an SNMP Host on Another Node

This procedure shows how to add an SNMP host on another node to the list of hosts that will receive trap notifications for the MIBs.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization.**

2   **Add the host to the SNMP host list of a community on another node.**

    phys-schost-1# **clsnmphost add -c** *SNMPcommunity* **-n** *node host*

    -c *SNMPcommunity*
      Specifies the SNMP community name that is used in conjunction with the host name.

      You must specify the SNMP community name *SNMPcommunity* when you add a host to a community other than `public`. If you use the `add` subcommand without the `-c` option, the subcommand uses `public` as the default community name.

      If the specified community name does not exist, this command creates the community.

    -n *node*
      Specifies the name of the *node* of the SNMP host that is provided access to the SNMP MIBs in the cluster. You can specify a node name or a node ID. If you do not specify this option, the current node is used by default.

    host
      Specifies the name, IP address, or IPv6 address of a host that is provided access to the SNMP MIBs in the cluster.

▼ How to Remove an SNMP Host on the Current Node

This procedure shows how to remove an SNMP host on the current node from the list of hosts that will receive trap notifications for the MIBs.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** `solaris.cluster.modify` **RBAC authorization.**

2   **Remove the host from the SNMP host list of a community on the current node.**

    phys-schost-1# **clsnmphost remove -c** *SNMPcommunity host*

    -c *SNMPcommunity*
      Specifies the name of the SNMP community from which the SNMP host is removed.

host

Specifies the name, IP address, or IPv6 address of the host that is removed from the configuration.

To remove all hosts in the specified SNMP community, use a plus sign (+) for *host* with the -c option. To remove all hosts, use the plus sign (+) for *host*.

▼ How to Remove an SNMP Host on Another Node

This procedure shows how to remove an SNMP host on another node from the list of hosts that will receive trap notifications for the MIBs.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization.**

2   **Remove the host from the SNMP host list of a community on the specified node.**

     `phys-schost-1# `**`clsnmphost remove -c`**` SNMPcommunity `**`-n`**` node  host`

remove

Removes the specified SNMP host from the specified node.

-c *SNMPcommunity*

Specifies the name of the SNMP community from which the SNMP host is removed.

-n *node*

Specifies the name of the *node* on which the SNMP host that is removed from the configuration. You can specify a node name or a node ID. If you do not specify this option, the current node is used by default.

host

Specifies the name, IP address, or IPv6 address of the host that is removed from the configuration.

To remove all hosts in the specified SNMP community, use a plus sign (+) for *host* with the -c option. To remove all hosts, use the plus sign (+) for *host*.

▼ How to Change an SNMP Event MIB

This procedure shows how to change the protocol for an SNMP event MIB.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Become superuser or assume a role that provides** solaris.cluster.modify **RBAC authorization.**

2   Change the protocol of the SNMP event MIB.

phys-schost-1# **clsnmpmib set -n** *node* **-p version=***value* *MIB*

-n *node*
    Specifies the *node* on which the event MIB that you want to change is located. You can specify a
    node ID or a node name. If you do not specify this option, the current node is used by default.

-p version=*value*
    Specifies the version of SNMP protocol to use with the MIBs. You specify *value* as follows:

- version=SNMPv2
- version=snmpv2
- version=2
- version=SNMPv3
- version=snmpv3
- version=3

MIB
    Specifies the name of the MIB or MIBs to which to apply the subcommand. If you do not specify
    this operand, the subcommand uses the default plus sign (+), which means all MIBs. If you use the
    MIB operand, specify the MIB in a space-delimited list after all other command-line options.

▼ How to Enable an SNMP Event MIB

This procedure shows how to enable an SNMP event MIB.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have
short forms. Except for the long and short forms of the command names, the commands are
identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization.**

2   Enable the SNMP event MIB.

phys-schost-1# **clsnmpmib enable -n** *node* *MIB*

-n *node*   Specifies the *node* on which the event MIB that you want to enable is located.
                   You can specify a node ID or a node name. If you do not specify this option,
                   the current node is used by default.

MIB   Specifies the name of the event MIB that you want to enable. You can use any
                   combination of uppercase and lowercase letters in *MIB*.

▼ How to Disable an SNMP Event MIB

This procedure shows how to disable an SNMP event MIB.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization.**

2 Disable the SNMP event MIB.

phys-schost-1# **clsnmpmib disable -n** *node MIB*

-n *node*   Specifies the *node* on which the event MIB that you want to disable is located. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

MIB   Specifies the name of the event MIB that you want to disable. You can use any combination of uppercase and lowercase letters in *MIB*.

▼ How to Add an SNMP User

This procedure shows how to add an SNMP user to the SNMP user configuration on a node.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization.**

2 Add the SNMP user.

phys-schost-1# **clsnmpuser create -n** *node* **-a** *authentication* \
                    -f *password user*

-n *node*   Specifies the node on which the SNMP user is added. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

-a *authentication*   Specifies the authentication protocol that is used to authorize the user. The value of the authentication protocol can be SHA or MD5.

-f *password*   Specifies a file that contains the SNMP user passwords. If you do not specify this option when you create a new user, the command prompts for a password. This option is valid only with the add subcommand.

                          You must specify user passwords on separate lines in the following format:

                          user:*password*

                          Passwords cannot contain the following characters or a space:

- ; (semicolon)
- : (colon)
- \ (backslash)
- \n (newline)

user   Specifies the name of the SNMP user that you want to add.

▼ How to Remove an SNMP User

This procedure shows how to remove an SNMP user from the SNMP user configuration on a node.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Become superuser or assume a role that provides `solaris.cluster.modify` **RBAC authorization.**

2   Remove the SNMP user.

phys-schost-1# **clsnmpuser delete -n** *node user*

-n *node*   Specifies the node from which the SNMP user is removed. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

user   Specifies the name of the SNMP user that you want to remove.

Configuring Control of CPU Usage

If you want to control the usage of CPU, configure the CPU control facility. For more information about configuring the CPU control facility, see the `rg_properties`(5) man page. This chapter provides information about the following topics:

Introduction to CPU Control

Sun Cluster enables you to control the usage of CPU. The configuration choices you can make on the Solaris 9 OS are not the same as the choices you can make on the Solaris 10 OS.

The CPU control facility builds on the functionality available in the Solaris OS. For information about zones, projects, resource pools, processor sets, and scheduling classes, see *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

SPARC: On the Solaris 9 OS, you can assign CPU shares to resource groups.

On the Solaris 10 OS, you can do the following:

- Assign CPU shares to resource groups.
- Assign processors to resource groups.

Note – All procedures in this chapter are for use on the Solaris 10 OS unless labeled as specific to the Solaris 9 OS.

Choosing a Scenario

Depending on the configuration choices you make and version of the operating system you choose, you can have different levels of CPU control. All aspects of CPU control described in this chapter are dependent on the resource group property `RG_SLM_TYPE` being set to `automated`.

Table 9–1 provides a description of the different configuration scenarios available.

TABLE 9–1 CPU Control Scenarios

Description	Instructions
SPARC: Resource group runs on the Solaris 9 OS Assign CPU shares to a resource group, providing a value for `project.cpu-shares`	"SPARC: How to Control CPU Usage on the Solaris 9 OS" on page 265
Resource group runs in the global zone on the Solaris 10 OS Assign CPU shares to resource groups and zones, providing values for `project.cpu-shares` and `zone.cpu-shares` You can perform this procedure whether or not non-global zones are configured.	"How to Control CPU Usage in the Global Zone" on page 266
Resource group runs in a non-global zone by using the default processor set Assign CPU shares to resource groups and zones, providing values for `project.cpu-shares` and `zone.cpu-shares` Perform this procedure if you do not need to control the size of the processor set.	"How to Control CPU Usage in a Non-Global Zone With the Default Processor Set" on page 268
Resource group runs in a non-global zone with a dedicated processor set Assign CPU shares to resource groups, providing values for `project.cpu-shares`, `zone.cpu-shares`, and maximum number of processors in a dedicated processor set Set the minimum number of processor sets in a dedicated processor set. Perform this procedure if you want to control CPU shares and the size of a processor set. You can exercise this control only in a non-global zone, by using a dedicated processor set.	"How to Control CPU Usage in a Non-Global Zone With a Dedicated Processor Set" on page 271

Fair Share Scheduler

The first step in the procedures to assign CPU shares to resource groups is to set the scheduler for the system to be the fair share scheduler (FSS). By default, the scheduling class for Solaris OS is timesharing schedule (TS). Set the scheduler to be FSS to have the shares configuration take effect.

You can create a dedicated processor set regardless of the scheduler class you choose.

Configuring CPU Control

The section includes the following procedures:

▼ SPARC: How to Control CPU Usage on the Solaris 9 OS

Perform this procedure to assign CPU shares to a resource group on a cluster running the Solaris 9 OS.

If a resource group is assigned CPU shares, Sun Cluster performs the following tasks when it starts a resource of the resource group:

- Creates a project named SCSLM_*resource_group_name* if that project does not already exist. This project is specific to the resource group and is assigned the specified number of CPU shares (project.cpu_shares).

- Starts the resource in the SCSLM_*resourcegroup_name* project.

For more information about configuring the CPU control facility, see the rg_properties(5) man page.

1   **Set the scheduler for the system to be the fair share scheduler (FSS).**

    # **dispadmin -d FSS**

    FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the priocntl command.

    # **priocntl -s -c FSS**

    Using the combination of the priocntl and dispadmin commands ensures that FSS becomes the default scheduler immediately and remains so after reboot. For more information about setting a scheduling class, see the dispadmin(1M) and priocntl(1) man pages.

    Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

    ---

2   **Configure the CPU control facility.**

    # **clresourcegroup create -p RG_SLM_TYPE=automated **
     [-p RG_SLM_CPU_SHARES=*value***]** *resource_group_name*

-p RG_SLM_TYPE=automated	Enables you to control CPU usage and automates some steps to configure the Solaris OS for system resource management.
-p RG_SLM_CPU-SHARES=*value*	Specifies the number of CPU shares assigned to the resource group-specific project, project.*cpu-shares*.
resource_group_name	Specifies the name of the resource group.

This step creates a resource group. You could alternatively use the clresourcegroup set command to modify an existing resource group.

3 Activate the configuration change.

```
# clresourcegroup online -M resource_group_name
```

resource_group_name   Specifies the name of the resource group.

Note – Do not remove or modify the SCSLM_*resource_group_name* project. You can add more resource control manually to the project, for example by configuring the project.max-lwps property. For more information, see the projmod(1M) man page.

▼ How to Control CPU Usage in the Global Zone

Perform this procedure to assign CPU shares to a resource group that will be executed in a global zone.

If a resource group is assigned CPU shares, Sun Cluster performs the following tasks when it starts a resource of the resource group in a global zone:

- Augments the number of CPU shares assigned to the global zone (zone.*cpu-shares*) with the specified number of CPU shares, if this has not already been done.
- Creates a project named SCSLM_*resourcegroup_name* in the global zone, if this has not already been done. This project is specific to the resource group, and is assigned the specified number of CPU shares (project.*cpu-shares*).
- Starts the resource in the SCSLM_*resourcegroup_name* project.

For more information about configuring the CPU control facility, see the rg_properties(5) man page.

1 Set the default scheduler for the system to be fair share scheduler (FSS).

```
# dispadmin -d FSS
```

FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the priocntl command.

```
# priocntl -s -C FSS
```

Using the combination of the priocntl and dispadmin commands ensures that FSS becomes the default scheduler immediately and remains so after reboot. For more information about setting a scheduling class, see the dispadmin(1M) and priocntl(1) man pages.

Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

---

2 On each node to use CPU control, configure the number of shares for the global zone and the minimum number of CPUs available in the default processor set.

Setting these parameters helps protect processes running in the global zone from competing for CPUs with processes running in non-global zones. If you do not assign a value to the globalzoneshares and defaultpsetmin properties, these properties take their default values.

```
# clnode set [-p globalzoneshares=integer] \
[-p defaultpsetmin=integer] \
node
```

-p defaultpsetmin=*defaultpsetmininteger*	Sets the minimum number of CPU available in the default processor set. The default value is 1.
-p globalzoneshares=*integer*	Sets the number of shares assigned to the global zone. The default value is 1.
node	Specifies nodes on which properties are to be set.

In setting these properties, you are setting properties for the global zone. If you do not set these properties, you cannot benefit from the RG_SLM_PSET_TYPE property in non-global zones.

3 Verify that you correctly set these properties.

```
# clnode show node
```

For the node you specify, the clnode command prints the properties set and the values that are set for these properties. If you do not set the CPU control properties with clnode, they take the default value.

4 Configure the CPU control facility.

```
# clresourcegroup create -p RG_SLM_TYPE=automated \
 [-p RG_SLM_CPU_SHARES=value] resource_group_name
```

-p RG_SLM_TYPE=automated	Enables you to control CPU usage and automates some steps to configure the Solaris OS for system resource management.
-p RG_SLM_CPU_SHARES=*value*	Specifies the number of CPU shares that are assigned to the resource group-specific project, project.cpu-shares and determines the number of CPU shares that are assigned to the global zone, zone.cpu-shares.
resource_group_name	Specifies the name of the resource group.

In this procedure, you do not set the RG_SLM_PSET_TYPE property. In the global zone, this property takes the value default.

This step creates a resource group. You could alternatively use the clresourcegroup set command to modify an existing resource group.

5 Activate the configuration change.

clresourcegroup online -M *resource_group_name*

resource_group_name   Specifies the name of the resource group.

Note – Do not remove or modify the SCSLM_*resource_group_name* project. You can add more resource control manually to the project, for example, by configuring the project.max-lwps property. For more information, see the projmod(1M) man page.

▼ How to Control CPU Usage in a Non-Global Zone With the Default Processor Set

Perform this procedure if you want to assign CPU shares for resource groups in a non-global zone, but do not need to create a dedicated processor set.

If a resource group is assigned CPU shares, Sun Cluster performs the following tasks when starting a resource of that resource group in a non-global zone:

- Creates a pool named SCSLM_*resource_group_name*, if this has not already been done.
- Associates the SCSLM_pool_*zone_name* pool to the default processor set.
- Dynamically binds the non-global zone to the SCSLM_pool*zone_name* pool.
- Augments the number of CPU shares assigned to the non-global zone (zone.cpu-shares) with the specified number of CPU shares, if this has not already been done.
- Creates a project named SCSLM_*resourcegroup_name* in the non-global zone, if this has not already been done. This project is specific to the resource group, and is assigned the specified number of CPU shares (*project.cpu-shares*).
- Starts the resource in the SCSLM_*resourcegroup_name* project.

For more information about configuring the CPU control facility, see the rg_properties(5) man page.

1 Set the default scheduler for the system to be fair share scheduler (FSS).

dispadmin -d FSS

FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the priocntl command:

```
# priocntl -s -C FSS
```

Using the combination of the priocntl and dispadmin commands ensures that FSS becomes the default schedule immediately and remains so after reboot. For more information about setting a scheduling class, see the dispadmin(1M) and priocntl(1) man pages.

Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

---

2 **On each node to use CPU control, configure the number of shares for the global zone and the minimum number of CPUs available in the default processor set.**

Setting these parameters helps protect processes running in the global zone from competing for CPUs with processes running in non-global zones. If you do not assign a value to the globalzoneshares and defaultpsetmin properties, these properties take their default values.

```
# clnode set [-p globalzoneshares=integer] \
[-p defaultpsetmin=integer] \
node
```

-p globalzoneshares=*integer*	Sets the number of shares assigned to the global zone. The default value is 1.
-p defaultpsetmin=*defaultpsetmininteger*	Sets the minimum number of CPUs available in the default processor set. The default value is 1.
node	Identifies nodes on which properties are to be set.

In setting these properties, you are setting properties for the global zone.

3 **Verify that you correctly set these properties:**

```
# clnode show node
```

For the node you specify, the clnode command prints the properties set and the values that are set for these properties. If you do not set the CPU control properties with clnode, they take the default value.

4 **Configure the CPU control facility.**

```
# clresourcegroup create -p RG_SLM_TYPE=automated \
 [-p RG_SLM_CPU_SHARES=value] resource_group_name
```

-p RG_SLM_TYPE=automated	Enables you to control CPU usage and automates some steps to configure the Solaris OS for system resource management.

-p RG_SLM_CPU_SHARES=*value*	Specifies the number of CPU shares assigned to the resource group-specific project (project.cpu-shares) and determines the number of CPU shares assigned to the non-global zone (zone.cpu_shares).
resource_group_name	Specifies the name of the resource group.

This step creates a resource group. You could alternatively use the clresourcegroup set command to modify an existing resource group.

You cannot set RG_SLM_TYPE to automated in a non-global zone if a pool other than the default pool is in the zone configuration or if the zone is dynamically bound to a pool other than the default pool. See the zonecfg(1M) and poolbind(1M) man pages for information about zone configuration and pool binding respectively. View your zone configuration as follows:

```
# zonecfg -z zone_name info pool
```

Note – A resource such as an HAStoragePlus or a LogicalHostname resource was configured to start in a non-global zone but with the GLOBAL_ZONE property set to TRUE is started in the global zone. Even if you set the RG_SLM_TYPE property to automated, this resource does not benefit from the CPU shares configuration and is treated as in a resource group with RG_SLM_TYPE set to manual.

In this procedure, you do not set the RG_SLM_PSET_TYPE property. Sun Cluster uses the default processor set.

5 Activate the configuration change.

```
# clresourcegroup online -M resource_group_name
```

resource_group_name   Specifies the name of the resource group.

If you set RG_SLM_PSET_TYPE to default, Sun Cluster creates a pool, SCSLM_pool_*zone_name*, but does not create a processor set. In this case, SCSLM_pool_*zone_name* is associated with the default processor set.

If online resource groups are no longer configured for CPU control in a non-global zone, the CPU share value for the non-global zone takes the value of zone.cpu-shares in the zone configuration. This parameter has a value of 1 by default. For more information about zone configuration, see the zonecfg(1M) man page.

Note – Do not remove or modify the SCSLM_*resource_group_name* project. You can add more resource control manually to the project, for example by configuring the project.max-lwps property. For more information, see the projmod(1M) man page.

▼ How to Control CPU Usage in a Non-Global Zone With a Dedicated Processor Set

Perform this procedure if you want your resource group to execute in a dedicated processor set.

If a resource group is configured to execute in a dedicated processor set, Sun Cluster performs the following tasks when it starts a resource of the resource group in a non-global zone:

- Creates a pool named SCSLM_pool_*zone_name*, if this has not already been done.
- Creates a dedicated processor set. The size of the processor set is determined by using the RG_SLM_CPU_SHARES and RG_SLM_PSET_MIN properties.
- Associates the SCSLM_pool_*zone_name* pool to the created processor set.
- Dynamically binds the non-global zone to the SCSLM_pool_*zone_name* pool.
- Augments the number of CPU shares assigned to the non-global zone with the specified number of CPU shares, if this has not already been done.
- Creates a project named SCSLM_*resourcegroup_name* in the non-global zone, if this has not already been done. This project is specific to the resource group, and is assigned the specified number of CPU shares (project.cpu-shares).
- Starts the resource in the SCSLM_*resourcegroup_name* project.

1 Set the scheduler for the system to be fair share scheduler (FSS).

```
# dispadmin -d FSS
```

FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the priocntl command.

```
# priocntl -s -C FSS
```

Using the combination of the priocntl and dispadmin commands ensures that FSS becomes the default schedule immediately and remains so after reboot. For more information about setting a scheduling class, see the dispadmin(1M) and priocntl(1) man pages.

Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

---

2 On each node to use CPU control, configure the number of shares for the global zone and the minimum number of CPUs available in the default processor set.

Setting these parameters helps protect processes running in the global zone from competing for CPUs with processes running in non-global zones. If you do not assign a value to the globalzoneshares and defaultpsetmin properties, these properties take their default values.

```
# clnode set  [-p globalzoneshares=integer] \
[-p defaultpsetmin=integer] \
node
```

-p defaultpsetmin=*defaultpsetmininteger*	Sets the minimum number of CPUs available in the default processor set. The default is 1.
-p globalzoneshares=*integer*	Sets the number of shares assigned to the global zone. The default is 1.
node	Identifies nodes on which properties are to be set.

In setting these properties, you are setting properties for the global zone.

3   Verify that you correctly set these properties:

clnode show *node*

For the node you specify, the clnode command prints the properties set and the values that are set for these properties. If you do not set the CPU control properties with clnode, they take the default value.

4   Configure the CPU control facility.

```
# clresourcegroup create -p RG_SLM_TYPE=automated \
 [-p RG_SLM_CPU_SHARES=value] \
-p -y RG_SLM_PSET_TYPE=value \
[-p RG_SLM_PSET_MIN=value] resource_group_name
```

-p RG_SLM_TYPE=automated	Enables you to control CPU control usage and automates some steps to configure the Solaris OS for system resource management.
-p RG_SLM_CPU_SHARES=*value*	Specifies the number of CPU shares assigned to the resource group-specific project (project.cpu-shares) and determines the number of CPU shares assigned to the non-global zone (zone.cpu-shares) and the maximum number of processors in a processor set.
-p RG_SLM_PSET_TYPE=*value*	Enables the creation of a dedicated processor set. To have a dedicated processor set, you can set this property to strong or weak. The values strong and weak are mutually exclusive. That is, you cannot configure resource groups in the same zone so that some are strong and others weak.
-p RG_SLM_PSET_MIN=*value*	Determines the minimum number of processors in the processor set.
resource_group_name	Specifies the name of the resource group.

This step creates a resource group. You could alternatively use the clresourcegroup set command to modify an existing resource group.

You cannot set RG_SLM_TYPE to automated in a non-global zone if a pool other than the default pool is in the zone configuration or if the zone is dynamically bound to a pool other than the default pool.

See the zonecfg(1M) and poolbind(1M) man pages for information about zone configuration and pool binding respectively. View your zone configuration as follows:

```
# zonecfg -z zone_name info pool
```

Note – A resource such as an HAStoragePlus or a LogicalHostname resource configured to start in a non-global zone but with the GLOBAL_ZONE property set to TRUE is started in the global zone. Even if you set the RG_SLM_TYPE property to automated, this resource does not benefit from the CPU shares and dedicated processor set configuration and is treated as in a resource group with RG_SLM_TYPE set to manual.

5   Activate the configuration change.

resource_group_name   Specifies the name of the resource group.

Note – Do not remove or modify the SCSLM_*resource_group_name* project. You can add more resource control manually to the project, for example by configuring the project.max-lwps property. For more information, see the projmod(1M) man page.

Changes made to RG_SLM_CPU_SHARES and RG_SLM_PSET_MIN while the resource group is online are taken into account dynamically. However, if RG_SLM_PSET_TYPE is set to strong, and if not enough CPUs are available to accommodate the change, the change requested for RG_SLM_PSET_MIN is not applied. In this case, a warning message is displayed. On next switchover, errors because insufficient CPUs can occur if not enough CPUs are available to acknowledge the values that you configured for RG_SLM_PSET_MIN.

If an online resource group is no longer configured for CPU control in the non-global zone, the CPU share value for the non-global zone takes the value of zone.cpu-shares. This parameter has a value of 1 by default.

10

Patching Sun Cluster Software and Firmware

This chapter provides the procedures for adding and removing patches for a Sun Cluster configuration that are contained in the following sections.

- "Overview of Patching Sun Cluster" on page 275
- "Patching Sun Cluster" on page 277

Overview of Patching Sun Cluster

Due to the nature of a cluster, all cluster member nodes must be at the same patch level for proper cluster operation. Occasionally, when patching a node with a Sun Cluster patch, you might need to temporarily remove a node from cluster membership or stop the entire cluster before installing the patch. This section describes these steps.

Before applying a Sun Cluster patch, check the Sun Cluster web site for any special instructions. For the current URL, see the *Sun Cluster 3.2 Release Notes for Solaris OS* or contact Enterprise Services. If no special instructions are provided, check the patch's README file.

Note – For Sun Cluster patches, always defer to the patch's README file and to SunSolve for instructions that supersede procedures in this chapter.

Patch installation on all cluster nodes can be described by one of the following scenarios:

Rebooting patch (node)   A node must be booted to single-user mode, using the
                                           command boot -sx or b -sx, before the patch can be
                                           applied, then rebooted to join the cluster. First you need
                                           to put the node into a "quiet" state by switching any
                                           resource groups or device groups from the node to be
                                           patched to another cluster member. Also, apply the
                                           patch to one cluster node at a time to avoid shutting
                                           down the entire cluster.

	The cluster itself remains available during this type of patch application, even though individual nodes are temporarily shut down. A patched node is able to rejoin a cluster as a member node even though other nodes are not yet at the same patch level.
Rebooting patch (cluster and firmware)	The cluster must be shut down and each node must be booted to single-user mode, using the command `boot -sx` or `b -sx`, to apply the software or firmware patch. Then, reboot the nodes to rejoin the cluster. For this type of patch, the cluster is unavailable during patch application.
Nonrebooting patch	A node does not have to be in a "quiet" state (it can still be mastering resource groups or device groups), nor does it have to be shut down or rebooted when applying the patch. However, you should still apply the patch to one node at a time and verify that the patch works before patching another node.

Note – Underlying cluster protocols do not change because of a patch.

Use the `patchadd` command to apply a patch to the cluster, and `patchrm` to remove a patch (when possible).

Sun Cluster Patch Tips

Use the following tips to help you administer Sun Cluster patches more efficiently:

- Refer to the Sun Cluster website for any special instructions associated with the patch or firmware update. For the current URL, see the *Sun Cluster 3.2 Release Notes for Solaris OS* or contact Enterprise Services.

- Always read the patch README file before applying the patch.

- Apply all patches (required and recommended) before running the cluster in a production environment.

- Check the hardware firmware levels and install any required firmware updates that might be needed.

- All nodes acting as cluster members must have the same patches.

- Keep cluster subsystem patches up to date. These patches include, for example, volume management, storage device firmware, and cluster transport.

- Review patch reports regularly, such as once a quarter, and patch a Sun Cluster configuration by using the recommended patch suite.

- Apply selective patches as recommended by Enterprise Services.
- Test failover after major patch updates. Be prepared to back out the patch if cluster operation is degraded or impaired.

Patching Sun Cluster

TABLE 10–1 Task Map: Patching the Cluster

Task	Instructions
Apply a nonrebooting Sun Cluster patch to one node at a time without stopping the node	"How to Apply a Nonrebooting Sun Cluster Patch" on page 284
Apply a rebooting Sun Cluster patch after taking the cluster member to noncluster mode	"How to Apply a Rebooting Patch (Node)" on page 277
	"How to Apply a Rebooting Patch (Cluster and Firmware)" on page 281
Remove a Sun Cluster patch	"Changing a Sun Cluster Patch" on page 285

▼ How to Apply a Rebooting Patch (Node)

Apply the patch to one node in the cluster at a time to keep the cluster itself operational during the patch process. With this procedure, you must first shut down the node and boot it to single-user mode by using the `boot -sx` or `b -sx` command, before applying the patch.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Before applying the patch, check the Sun Cluster product web site for any special preinstallation or postinstallation instructions.**

2   **Become superuser or assume a role that provides** `solaris.cluster.admin` **RBAC authorization on the node to which you are applying the patch.**

3   **List the resource groups and device groups on the node being patched.**

```
# clresourcegroup status -n node
# cldevicegroup status -n node
```

4   **Switch all resource groups, resources, and device groups from the node being patched to other cluster members.**

```
# clnode evacuate -n node
```

evacuate　　　Evacuates all device groups and resource groups including all non-global zones from
　　　　　　　the specified node.

-n *node*　　　Specifies the node from which you are switching the resource groups and device
　　　　　　　groups.

5　Shut down the node.

```
# shutdown -g0 [-y]
 [-i0]
```

6　Boot the node in noncluster, single-user mode.

- On SPARC based systems, perform the following command:

  ```
  ok boot -sx
  ```

- On x86 based system, perform the following commands:

  ```
  phys-schost# shutdown -g -y -i0

  Press any key to continue
  ```

  a.　In the GRUB menu, use the arrow keys to select the appropriate Solaris entry and type e **to edit its commands.**

  The GRUB menu appears similar to the following:

  ```
  GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
  +-----------------------------------------------------------------------+
  | Solaris 10 /sol_10_x86                                                 |
  | Solaris failsafe                                                       |
  |                                                                       |
  +-----------------------------------------------------------------------+
  Use the ^ and v keys to select which entry is highlighted.
  Press enter to boot the selected OS, 'e' to edit the
  commands before booting, or 'c' for a command-line.
  ```

  For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration* in *System Administration Guide: Basic Administration*.

  b.　In the boot parameters screen, use the arrow keys to select the kernel entry and type e **to edit the entry.**

  The GRUB boot parameters screen appears similar to the following:

  ```
  GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
  +-----------------------------------------------------------------------+
  | root (hd0,0,a)                                                         |
  | kernel /platform/i86pc/multiboot                                      |
  | module /platform/i86pc/boot_archive                                   |
  ```

```
+---------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

c. Add -x **to the command to specify that the system boot into noncluster mode.**

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. ]

grub edit> kernel /platform/i86pc/multiboot -x
```

d. Press the Enter key to accept the change and return to the boot parameters screen.

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+---------------------------------------------------------------------+
| root (hd0,0,a)                                                      |
| kernel /platform/i86pc/multiboot -x                                 |
| module /platform/i86pc/boot_archive                                 |
+---------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.-
```

e. Type b **to boot the node into noncluster mode.**

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -x option to the kernel boot parameter command.

---

7 Apply the patch.

```
# patchadd -M patch-dir patch-id
```

patch-dir   Specifies the directory location of the patch.

patch-id   Specifies the patch number of a given patch.

Note – Always defer to the instructions in the patch directory, which supersede procedures in this chapter.

---

8 Verify that the patch has been installed successfully.

showrev -p | grep *patch-id*

9 Reboot the node into the cluster.

reboot

10 Verify that the patch works, and that the node and cluster are operating normally.

11 Repeat Step 2 through Step 10 for all remaining cluster nodes.

12 Switch resource groups and device groups as needed.

After you reboot all the nodes, the last node rebooted will not have the resource groups and device groups online.

cldevicegroup switch -n *node*   + | *devicegroup* ...
clresourcegroup switch -n *node*[:*zone*][*,...*] + | *resource-group* ...

node   The name of the node to which you are switching the resource groups and device groups.

zone   The name of the non-global zone on node that can master the resource group. Specify zone only if you specified a non-global zone when you created the resource group.

Example 10–1 Applying a Rebooting Patch (Node)

The following example shows the application of a rebooting Sun Cluster patch to a node.

```
# clresourcegroup status -n rg1
...Resource Group     Resource
--------------     --------
rg1                rs-2
rg1                rs-3
...
# cldevicegroup status -n nodedg-schost-1
...
Device Group Name:                                    dg-schost-1
...
# clnode evacuate  phys-schost-2
# shutdown -g0 -y -i0
...
```

Boot the node in noncluster, single-user mode.

- SPARC: Type:

    ok **boot -sx**

- x86: Boot the node in noncluster, single-user mode. See boot steps in the proceeding procedure.

```
# patchadd -M /var/tmp/patches 234567-05
...
# showrev -p | grep 234567-05

...
# reboot
...
# cldevicegroup switch -n  phys-schost-1 dg-schost-1
# clresourcegroup switch -n  phys-schost-1 schost-sa-1
```

See Also   If you need to back out a patch, see "Changing a Sun Cluster Patch" on page 285.

▼ How to Apply a Rebooting Patch (Cluster and Firmware)

With this procedure, you must first shut down the cluster and boot each node to single-user mode by using the boot -sx or b -sx command, before applying the patch.

1   **Before applying the patch, check the Sun Cluster product web site for any special preinstallation or postinstallation instructions.**

2   **Become superuser on any node in the cluster.**

3   **Shut down the cluster.**

cluster shutdown -y -g *grace-period* **"***message***"**

-y   Specifies to answer *yes* to the confirmation prompt.

-g *grace-period*   Specifies, in seconds, the amount of time to wait before shutting down. Default grace period is 60 seconds.

message   Specifies the warning message to broadcast. Use quotes if *message* contains multiple words.

4   **Boot each node into noncluster, single-user mode.**

On the console of each node, run the following command.

- On SPARC based systems, perform the following command:

    ok **boot -sx**

- On x86 based system, perform the following commands:.

```
phys-schost# shutdown -g -y -i0

Press any key to continue
```

a. In the GRUB menu, use the arrow keys to select the appropriate Solaris entry and type e to edit its commands.

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----------------------------------------------------------------------+
| Solaris 10 /sol_10_x86                                                 |
| Solaris failsafe                                                       |
|                                                                        |
+-----------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration* in *System Administration Guide: Basic Administration*.

b. In the boot parameters screen, use the arrow keys to select the kernel entry and type e to edit the entry.

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----------------------------------------------------------------------+
| root (hd0,0,a)                                                         |
| kernel /platform/i86pc/multiboot                                       |
| module /platform/i86pc/boot_archive                                    |
+-----------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

c. Add -x to the command to specify that the system boot into noncluster mode.

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. ]

grub edit> kernel /platform/i86pc/multiboot -x
```

    d. Press the Enter key to accept the change and return to the boot parameters screen.

      The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+----------------------------------------------------------------------+
| root (hd0,0,a)                                                       |
| kernel /platform/i86pc/multiboot -x                                 |
| module /platform/i86pc/boot_archive                                 |
+----------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.-
```

    e. Type b **to boot the node into noncluster mode.**

    Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -x option to the kernel boot parameter command.

---

5 Apply the software or firmware patch.

On one node at a time, run the following command.

`# patchadd -M `*patch-dir patch-id*

patch-dir   Specifies the directory location of the patch.

patch-id   Specifies the patch number of a given patch.

Note – Always defer to the instructions in the patch directory that supersede procedures in this chapter.

---

6 Verify that the patch has been installed successfully on each node.

`# showrev -p | grep `*patch-id*

7 After applying the patch to all nodes, reboot the nodes into the cluster.

On each node, run the following command.

`# reboot`

8 Verify that the patch works, and that the nodes and cluster are operating normally.

Example 10–2   Applying a Rebooting Patch (Cluster)

The following example shows the application of a rebooting Sun Cluster patch to a cluster.

```
# cluster shutdown -g0 -y
...
```

Boot the cluster in noncluster, single-user mode.

- SPARC: Type:

  ```
  ok boot -sx
  ```

- x86: Boot each node into noncluster, single-user mode. See the proceeding procedure for steps.

```
...
# patchadd -M /var/tmp/patches 234567-05
```
(*Apply patch to other cluster nodes*)
```
...
# showrev -p | grep 234567-05
# reboot
```

See Also   If you need to back out a patch, see "Changing a Sun Cluster Patch" on page 285.

▼ How to Apply a Nonrebooting Sun Cluster Patch

Apply the patch to one node in the cluster at a time. When applying a nonrebooting patch, you do not need to first shut down the node that is receiving the patch.

1 **Before applying the patch, check the Sun Cluster product web page for any special preinstallation or postinstallation instructions.**

2 **Apply the patch on a single node.**

   ```
   # patchadd -M patch-dir patch-id
   ```

   patch-dir   Specifies the directory location of the patch.

   patch-id   Specifies the patch number of a given patch.

3 **Verify that the patch has been installed successfully.**

   ```
   # showrev -p | grep patch-id
   ```

4 **Verify that the patch works, and that the node and cluster are operating normally.**

5 **Repeat Step 2 through Step 4 for the remaining cluster nodes.**

Example 10–3 Applying a Nonrebooting Sun Cluster Patch

```
# patchadd -M /tmp/patches 234567-05
...
# showrev -p | grep 234567-05
```

See Also If you need to back out a patch, see "Changing a Sun Cluster Patch" on page 285.

Changing a Sun Cluster Patch

To remove a Sun Cluster patch, you must re-apply the previous patch or update release by following. To re-apply a previous Sun Cluster patch, see "How to Apply a Rebooting Patch (Cluster and Firmware)" on page 281.

◆ ◆ ◆ **CHAPTER 11**

11

Backing Up and Restoring a Cluster

This chapter provides the following sections:

-
-

Backing Up a Cluster

TABLE 11–1 Task Map: Backing Up Cluster Files

Task	Instructions
Find the names of the file systems you want to back up	"How to Find File-System Names to Back Up" on page 288
Calculate how many tapes you need to contain a full backup	"How to Determine the Number of Tapes Needed for a Full Backup" on page 288
Back up the root file system	"How to Back Up the Root (/) File System" on page 289
Perform online backup for mirrored or plexed file systems	"How to Perform Online Backups for Mirrors (Solaris Volume Manager)" on page 290
	"SPARC: How to Perform Online Backups for Volumes (VERITAS Volume Manager)" on page 293
Back up the cluster configuration	"How to Back Up the Cluster Configuration" on page 297
Back up disk partitioning configuration for storage disk	See the documentation for your storage disk

▼ How to Find File-System Names to Back Up

Use this procedure to determine the names of the file systems that you want to back up.

1 Display the contents of the `/etc/vfstab` **file.**

You do not need to be superuser or assume an equivalent role to run this command.

```
% more /etc/vfstab
```

2 Look in the mount-point column for the name of the file system that you are backing up.

Use this name when you back up the file system.

```
% more /etc/vfstab
```

Example 11–1   Finding File System Names to Back Up

The following example displays the names of available file systems that are listed in the `/etc/vfstab` file.

```
% more /etc/vfstab
#device              device           mount FS fsck  mount  mount
#to mount            to fsck          point type  pass  at boot options
#
#/dev/dsk/c1d0s2     /dev/rdsk/c1d0s2 /usr   ufs   1     yes    -
 f                   -                /dev/fd fd   -     no     -
 /proc               -                /proc  proc  -     no     -
 /dev/dsk/c1t6d0s1   -                -      swap  -     no     -
 /dev/dsk/c1t6d0s0   /dev/rdsk/c1t6d0s0 /    ufs   1     no     -
 /dev/dsk/c1t6d0s3   /dev/rdsk/c1t6d0s3 /cache ufs 2     yes    -
 swap                -                /tmp   tmpfs -     yes    -
```

▼ How to Determine the Number of Tapes Needed for a Full Backup

Use this procedure to calculate the number of tapes that you need to back up a file system.

1 Become superuser or assume an equivalent role on the cluster node that you are backing up.

2 Estimate the size of the backup in bytes.

```
# ufsdump S filesystem
```

S   Displays the estimated number of bytes needed to perform the backup.

filesystem   Specifies the name of the file system you want to back up.

3 Divide the estimated size by the capacity of the tape to see how many tapes you need.

Example 11–2   Determining the Number of Tapes Needed

In the following example, the file system size of 905,881,620 bytes easily fits on a 4-Gbyte tape
(905,881,620 ÷ 4,000,000,000).

```
# ufsdump S /global/phys-schost-1
905881620
```

▼ How to Back Up the Root (/) File System

Use this procedure to back up the root (/) file system of a cluster node. Ensure that the cluster is
running without errors before performing the backup procedure.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have
short forms. Except for the long and short forms of the command names, the commands are
identical. For a list of the commands and their short forms, see Appendix A.

1. **Become superuser or assume a role the provides** `solaris.cluster.modify` **RBAC authorization on the cluster node that you are backing up.**

2. **Switch each running data service from the node to be backed up to another node in the cluster.**

    ```
    # clnode evacuate node
    ```

    node   Specifies the node from which you are switching resource groups and device groups.

3. **Shut down the node.**

    ```
    # shutdown -g0 -y -i0
    ```

4. **Reboot the node in noncluster mode.**

    - SPARC: Type:

        ```
        ok boot -x
        ```

    - x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. On the Current Boot Parameters screen, type b or i.

        ```
                            <<< Current Boot Parameters >>>
        Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
        sd@0,0:a
        Boot args:

        Type  b [file-name] [boot-flags] <ENTER> to boot with options
        or    i <ENTER>                          to enter boot interpreter
        or      <ENTER>                          to boot with defaults
        ```

```
                        <<< timeout in 5 seconds >>>
        Select (b)oot or (i)nterpreter: b -x
```

5   Back up the root (/) file system.

■ If the root disk is not encapsulated, use the following command.

   # **ufsdump 0ucf** *dump-device* **/**

■ If the root disk is encapsulated, use the following command.

   # **ufsdump 0ucf** *dump-device* **/dev/vx/rdsk/rootvol**

Refer to the ufsdump(1M) man page for more information.

6   Reboot the node in cluster mode.

   # **init 6**

Example 11–3   Backing Up the Root (/) File System

In the following example, the root (/) file system is backed up onto tape device /dev/rmt/0.

```
# ufsdump 0ucf /dev/rmt/0 /
  DUMP: Writing 63 Kilobyte records
  DUMP: Date of this level 0 dump: Tue Apr 18 18:06:15 2000
  DUMP: Date of last level 0 dump: the epoch
  DUMP: Dumping /dev/rdsk/c0t0d0s0 (phys-schost-1:/) to /dev/rmt/0
  DUMP: Mapping (Pass I) [regular files]
  DUMP: Mapping (Pass II) [directories]
  DUMP: Estimated 859086 blocks (419.48MB).
  DUMP: Dumping (Pass III) [directories]
  DUMP: Dumping (Pass IV) [regular files]
  DUMP: 859066 blocks (419.47MB) on 1 volume at 2495 KB/sec
  DUMP: DUMP IS DONE
  DUMP: Level 0 dump on Tue Apr 18 18:06:15 2000
```

▼ How to Perform Online Backups for Mirrors (Solaris Volume Manager)

A mirrored Solstice DiskSuite metadevice or Solaris Volume Manager volume can be backed up without unmounting it or taking the entire mirror offline. One of the submirrors must be taken offline temporarily, thus losing mirroring, but it can be placed online and resynchronized as soon as the backup is complete, without halting the system or denying user access to the data. Using mirrors to perform online backups creates a backup that is a "snapshot" of an active file system.

A problem might occur if a program writes data onto the volume immediately before the lockfs command is run. To prevent this problem, temporarily stop all the services running on this node. Also, ensure the cluster is running without errors before performing the backup procedure.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume an equivalent role on the cluster node that you are backing up.**

2 **Use the** metaset**(1M) command to determine which node has the ownership on the backed-up volume.**

    # **metaset -s** *setname*

    -s *setname*   Specifies the disk set name.

3 **Use the** lockfs**(1M) command with the** -w **option to lock the file system from writes.**

    # **lockfs -w** *mountpoint*

Note – You must lock the file system only if a UFS file system resides on the mirror. For example, if the Solstice DiskSuite metadevice or Solaris Volume Manager volume is set up as a raw device for database management software or some other specific application, you do not need to use the lockfs command. You might, however, run the appropriate vendor-dependent utility to flush any buffers and lock access.

---

4 **Use the** metastat**(1M) command to determine the names of the submirrors.**

    # **metastat -s** *setname* **-p**

    -p   Displays the status in a format similar to the md.tab file.

5 **Use the** metadetach**(1M) command to take one submirror offline from the mirror.**

    # **metadetach -s** *setname* *mirror* *submirror*

Note – Reads continue to be made from the other submirrors. However, the offline submirror is unsynchronized as soon as the first write is made to the mirror. This inconsistency is corrected when the offline submirror is brought back online. You do not need to run fsck.

---

6 **Unlock the file systems and allow writes to continue, using the** lockfs **command with the** -u **option.**

    # **lockfs -u** *mountpoint*

7 **Perform a file-system check.**

    # **fsck /dev/md/***diskset*/rdsk/*submirror*

8 Back up the offline submirror to tape or another medium.

Use the ufsdump(1M) command or the backup utility that you usually use.

ufsdump 0ucf *dump-device submirror*

Note – Use the raw device (/rdsk) name for the submirror, rather than the block device (/dsk) name.

---

9 Use the metattach**(1M) command to place the metadevice or volume back online.**

metattach -s *setname mirror submirror*

When the metadevice or volume is placed online, it is automatically resynchronized with the mirror.

10 Use the metastat **command to verify that the submirror is resynchronizing.**

metastat -s *setname mirror*

Example 11–4 Performing Online Backups for Mirrors (Solaris Volume Manager)

In the following example, the cluster node phys-schost-1 is the owner of the metaset schost-1, therefore the backup procedure is performed from phys-schost-1. The mirror /dev/md/schost-1/dsk/d0 consists of the submirrors d10, d20, and d30.

```
[Determine the owner of the metaset:]
# metaset -s schost-1
Set name = schost-1, Set number = 1
Host              Owner
  phys-schost-1   Yes
...
[Lock the file system from writes:]
# lockfs -w /global/schost-1
[List the submirrors:]
# metastat -s schost-1 -p
schost-1/d0 -m schost-1/d10 schost-1/d20 schost-1/d30 1
schost-1/d10 1 1 d4s0
schost-1/d20 1 1 d6s0
schost-1/d30 1 1 d8s0
[Take a submirror offline:]
# metadetach -s schost-1 d0 d30
[Unlock the file system:]
# lockfs -u /
[Check the file system:]
# fsck /dev/md/schost-1/rdsk/d30
[Copy the submirror to the backup device:]
# ufsdump 0ucf /dev/rmt/0 /dev/md/schost-1/rdsk/d30
  DUMP: Writing 63 Kilobyte records
  DUMP: Date of this level 0 dump: Tue Apr 25 16:15:51 2000
  DUMP: Date of last level 0 dump: the epoch
```

```
  DUMP: Dumping /dev/md/schost-1/rdsk/d30 to /dev/rdsk/c1t9d0s0.
  ...
  DUMP: DUMP IS DONE
```
[*Bring the submirror back online:*]
 # metattach -s schost-1 d0 d30
```
schost-1/d0: submirror schost-1/d30 is attached
```
[*Resynchronize the submirror:*]
 # metastat -s schost-1 d0
```
schost-1/d0: Mirror
    Submirror 0: schost-0/d10
      State: Okay
    Submirror 1: schost-0/d20
      State: Okay
    Submirror 2: schost-0/d30
      State: Resyncing
    Resync in progress: 42% done
    Pass: 1
    Read option: roundrobin (default)
...
```

▼ SPARC: How to Perform Online Backups for Volumes (VERITAS Volume Manager)

VERITAS Volume Manager identifies a mirrored volume as a plex. A plex can be backed up without unmounting it or taking the entire volume offline. This result is accomplished by creating a snapshot copy of the volume and backing up this temporary volume without halting the system or denying user access to the data.

Ensure that the cluster is running without errors before performing the backup procedure.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   **Log on to any node in the cluster, and become superuser or assume a role that provides** solaris.cluster.admin **RBAC authorization on the current primary node for the disk group on the cluster.**

2   **List the disk group information.**

    # vxprint -g *diskgroup*

3   **Determine which node has the disk group currently imported, indicating it is the primary node for the disk group.**

    # cldevicegroup status

4   **Create a snapshot of the volume.**

    # **vxassist -g** *diskgroup* **snapstart** *volume*

Note – Creating a snapshot can take a long time, depending on the size of your volume.

---

5   **Verify that the new volume was created.**

    # **vxprint -g** *diskgroup*

    When the snapshot is complete, a status of Snapdone displays in the State field for the selected disk
    group.

6   **Stop any data services that are accessing the file system.**

    # **clresourcegroup offline** *resource-group*

Note – Stop all data services to ensure that the data file system is properly backed up. If no data
services are running, you do not need to perform Step 6 and Step 8.

---

7   **Create a backup volume named** bkup-vol **and attach the snapshot volume to it.**

    # **vxassist -g** *diskgroup* **snapshot** *volume* **bkup-vol**

8   **Restart any data services that were stopped in Step 6, using the** clresourcegroup **command.**

    # **clresourcegroup online** *- zone* **-n** *node resourcegroup*

    node   The name of the node.

    zone   The name of the non-global zone on *node* that can master the resource group. Specify *zone*
             only if you specified a non-global zone when you created the resource group.

9   **Verify the volume is now attached to the new volume** bkup-vol**.**

    # **vxprint -g** *diskgroup*

10 **Register the device group configuration change.**

    # **cldevicegroup sync** *diskgroup*

11 **Check the backup volume.**

    # **fsck -y /dev/vx/rdsk/***diskgroup*/*bkup-vol*

12 **Perform a backup to copy the volume** bkup-vol **to tape or another medium.**

    Use the ufsdump(1M) command or the backup utility that you normally use.

    # **ufsdump 0ucf** *dump-device* **/dev/vx/dsk/***diskgroup*/*bkup-vol*

13 **Remove the temporary volume.**

vxedit -rf rm *bkup-vol*

14 **Register the disk group configuration changes.**

cldevicegroup sync *diskgroup*

Example 11–5 SPARC: Performing Online Backups for Volumes (VERITAS Volume Manager)

In the following example, the cluster node phys-schost-2 is the primary owner of the device group schost-1. Therefore, the backup procedure is performed from phys-schost-2. The volume /vol01 is copied and then associated with a new volume, bkup-vol.

```
[Become superuser or assume a role that provides solaris.cluster.admin RBAC authorization on the primary node.]
[Identify the current primary node for the device group:]
# cldevicegroup status
-- Device Group Servers --
                        Device Group     Primary          Secondary
                        ------------     -------          ---------
 Device group servers:  rmt/1            -                -
 Device group servers:  schost-1         phys-schost-2    phys-schost-1

-- Device Group Status --
                        Device Group        Status
                        ------------        ------
 Device group status:   rmt/1               Offline
 Device group status:   schost-1            Online
[List the device group information:]
# vxprint -g schost-1
TY NAME          ASSOC      KSTATE    LENGTH    PLOFFS STATE    TUTIL0  PUTIL0
dg schost-1      schost-1   -         -         -      -        -       -

dm schost-101    c1t1d0s2   -         17678493  -      -        -       -
dm schost-102    c1t2d0s2   -         17678493  -      -        -       -
dm schost-103    c2t1d0s2   -         8378640   -      -        -       -
dm schost-104    c2t2d0s2   -         17678493  -      -        -       -
dm schost-105    c1t3d0s2   -         17678493  -      -        -       -
dm schost-106    c2t3d0s2   -         17678493  -      -        -       -

v  vol01         gen        ENABLED   204800    -      ACTIVE   -       -
pl vol01-01      vol01      ENABLED   208331    -      ACTIVE   -       -
sd schost-101-01 vol01-01   ENABLED   104139    0      -        -       -
sd schost-102-01 vol01-01   ENABLED   104139    0      -        -       -
pl vol01-02      vol01      ENABLED   208331    -      ACTIVE   -       -
sd schost-103-01 vol01-02   ENABLED   103680    0      -        -       -
sd schost-104-01 vol01-02   ENABLED   104139    0      -        -       -
pl vol01-03      vol01      ENABLED   LOGONLY   -      ACTIVE   -       -
```

```
sd schost-103-02  vol01-03   ENABLED  5         LOG    -        -       -
```
[*Start the snapshot operation*:]
```
# vxassist -g schost-1 snapstart vol01
```
[*Verify the new volume was created*:]
```
# vxprint -g schost-1
TY NAME           ASSOC    KSTATE    LENGTH   PLOFFS STATE    TUTIL0  PUTIL0
dg schost-1       schost-1    -       -        -      -        -       -

dm schost-101     c1t1d0s2    -       17678493 -      -        -       -
dm schost-102     c1t2d0s2    -       17678493 -      -        -       -
dm schost-103     c2t1d0s2    -       8378640  -      -        -       -
dm schost-104     c2t2d0s2    -       17678493 -      -        -       -
dm schost-105     c1t3d0s2    -       17678493 -      -        -       -
dm schost-106     c2t3d0s2    -       17678493 -      -        -       -

v  vol01          gen      ENABLED   204800   -      ACTIVE   -       -
pl vol01-01       vol01    ENABLED   208331   -      ACTIVE   -       -
sd schost-101-01  vol01-01 ENABLED   104139   0      -        -       -
sd schost-102-01  vol01-01 ENABLED   104139   0      -        -       -
pl vol01-02       vol01    ENABLED   208331   -      ACTIVE   -       -
sd schost-103-01  vol01-02 ENABLED   103680   0      -        -       -
sd schost-104-01  vol01-02 ENABLED   104139   0      -        -       -
pl vol01-03       vol01    ENABLED   LOGONLY  -      ACTIVE   -       -
sd schost-103-02  vol01-03 ENABLED   5        LOG    -        -       -
pl vol01-04       vol01    ENABLED   208331   -      SNAPDONE -       -
sd schost-105-01  vol01-04 ENABLED   104139   0      -        -       -
sd schost-106-01  vol01-04 ENABLED   104139   0      -        -       -
```
[*Stop data services, if necessary*:]
```
# clresourcegroup offline nfs-rg
```
[Create a copy of the volume:]
```
# vxassist -g schost-1 snapshot vol01 bkup-vol
```
[*Restart data services, if necessary*:]
```
# clresourcegroup online -n phys-schost-1 nfs-rg
```
[*Verify bkup-vol was created*:]
```
# vxprint -g schost-1
TY NAME           ASSOC       KSTATE    LENGTH   PLOFFS STATE    TUTIL0  PUTIL0
dg schost-1       schost-1      -       -        -      -        -       -

dm schost-101     c1t1d0s2      -       17678493 -      -        -       -
...

v  bkup-vol       gen         ENABLED   204800   -      ACTIVE   -       -
pl bkup-vol-01    bkup-vol    ENABLED   208331   -      ACTIVE   -       -
sd schost-105-01  bkup-vol-01 ENABLED   104139   0      -        -       -
sd schost-106-01  bkup-vol-01 ENABLED   104139   0      -        -       -

v  vol01          gen         ENABLED   204800   -      ACTIVE   -       -
```

```
pl vol01-01       vol01        ENABLED  208331   -      ACTIVE   -      -
sd schost-101-01  vol01-01     ENABLED  104139   0      -        -      -
sd schost-102-01  vol01-01     ENABLED  104139   0      -        -      -
pl vol01-02       vol01        ENABLED  208331   -      ACTIVE   -      -
sd schost-103-01  vol01-02     ENABLED  103680   0      -        -      -
sd schost-104-01  vol01-02     ENABLED  104139   0      -        -      -
pl vol01-03       vol01        ENABLED  LOGONLY  -      ACTIVE   -      -
sd schost-103-02  vol01-03     ENABLED  5        LOG    -        -      -
```
[*Synchronize the disk group with cluster framework:*]
 # cldevicegroup sync schost-1
[Check the file systems:]
 # fsck -y /dev/vx/rdsk/schost-1/bkup-vol
[*Copy bkup-vol to the backup device*:]
 # ufsdump 0ucf /dev/rmt/0 /dev/vx/rdsk/schost-1/bkup-vol
```
  DUMP: Writing 63 Kilobyte records
  DUMP: Date of this level 0 dump: Tue Apr 25 16:15:51 2000
  DUMP: Date of last level 0 dump: the epoch
  DUMP: Dumping /dev/vx/dsk/schost-2/bkup-vol to /dev/rmt/0.
  ...
  DUMP: DUMP IS DONE
```
[*Remove the bkup-volume*:]
 # vxedit -rf rm bkup-vol
[*Synchronize the disk group*:]
 # cldevicegroup sync schost-1

▼ How to Back Up the Cluster Configuration

To ensure that your cluster configuration is archived and to facilitate easy recovery of the your cluster configuration, periodically back up your cluster configuration. Sun Cluster 3.2 provides the ability to export your cluster configuration to an eXtensible Markup Language (XML) file.

1   Log on to any node in the cluster, and become superuser or assume a role that provides solaris.cluster.read **RBAC authorization.**

2   Export the cluster configuration information to a file.

 # /usr/cluster/bin/cluster/export -o *configfile*

configfile   The name of the XML configuration file that the cluster command is exporting the cluster configuration information to. For information on the XML configuration file, see clconfiguration(5CL)

3   Verify that the cluster configuration information was successfully exported to the XML file.

 # vi *configfile*

Restoring Cluster Files

The ufsrestore(1M) command copies files to disk, relative to the current working directory, from backups created by using the ufsdump(1M) command. You can use ufsrestore to reload an entire file system hierarchy from a level 0 dump and incremental dumps that follow it, or to restore one or more single files from any dump tape. If ufsrestore is run as superuser or assumed an equivalent role, files are restored with their original owner, last modification time, and mode (permissions).

Before you start to restore files or file systems, you need to know the following information.

- Which tapes you need
- The raw device name on which you are restoring the file system
- The type of tape drive you are using
- The device name (local or remote) for the tape drive
- The partition scheme of any failed disk, because the partitions and file systems must be exactly duplicated on the replacement disk

TABLE 11–2 Task Map: Restoring Cluster Files

Task	Instructions
For Solaris Volume Manager, restore files interactively	"How to Restore Individual Files Interactively (Solaris Volume Manager)" on page 298
For Solaris Volume Manager, restore the root (/) file system	"How to Restore the Root (/) File System (Solaris Volume Manager)" on page 299
	"How to Restore a Root (/) File System That Was on a Solstice DiskSuite Metadevice or Solaris Volume Manager Volume" on page 301
For VERITAS Volume Manager, restore a non-encapsulated root (/) file system	"SPARC: How to Restore a Non-Encapsulated Root (/) File System (VERITAS Volume Manager)" on page 306
For VERITAS Volume Manager, restore an encapsulated root (/) file system	"SPARC: How to Restore an Encapsulated Root (/) File System (VERITAS Volume Manager)" on page 308

▼ How to Restore Individual Files Interactively (Solaris Volume Manager)

Use this procedure to restore one or more individual files. Ensure that the cluster is running without errors before performing the restore procedure.

1   Become superuser or assume a role that provides solaris.cluster.admin **RBAC authorization on the cluster node you are restoring.**

2 Stop all the data services that are using the files to be restored.

```
# clresourcegroup offline resource-group
```

3 Restore the files.

```
# ufsrestore
```

▼ How to Restore the Root (/) File System (Solaris Volume Manager)

Use this procedure to restore the root (/) file systems to a new disk, such as after replacing a bad root disk. The node being restored should not be booted. Ensure that the cluster is running without errors before performing the restore procedure.

Note – Because you must partition the new disk by using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and re-create file systems as appropriate.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 Become superuser or assume a role that provides solaris.cluster.modify **RBAC authorization on a cluster node with access to the disk sets to which the node to be restored is also attached.**

Use a node *other than* the node that you are restoring.

2 Remove the host name of the node being restored from all metasets.

Run this command from a node in the metaset other than the node that you are removing. Because the recovering node is offline, the system will display an RPC: Rpcbind failure - RPC: Timed out error. Ignore this error and continue to the next step.

```
# metaset -s setname -f -d -h nodelist
```

-s setname	Specifies the disk set name.
-f	Deletes the last host from the diskset.
-d	Deletes from the disk set.
-h *nodelist*	Specifies the name of the node to delete from the disk set.

3 Restore the root (/) and /usr **file systems.**

To restore the root and /usr file systems, follow the procedure in Chapter 27, "Restoring Files and File Systems (Tasks)," in *System Administration Guide: Devices and File Systems*. Omit the step in the Solaris OS procedure to reboot the system.

> **Note** – Ensure that you create the `/global/.devices/node@nodeid` file system.

---

4. **Reboot the node in multiuser mode.**

   ```
   # reboot
   ```

5. **Replace the disk ID.**

   ```
   # cldevice repair rootdisk
   ```

6. **Use the** metadb**(1M) command to re-create the state database replicas.**

   ```
   # metadb -c copies -af raw-disk-device
   ```

-c *copies*	Specifies the number of replicas to create.
-f *raw-disk-device*	Raw disk device on which to create replicas.
-a	Adds replicas.

7. **From a cluster node other than the restored node add the restored node to all disk sets.**

   ```
   phys-schost-2# metaset -s setname -a -h nodelist
   ```

-a	Creates and adds the host to the disk set.

   The node is rebooted into cluster mode. The cluster is ready to use.

Example 11–6   Restoring the Root (/) File System (Solaris Volume Manager)

The following example shows the root (/) file system restored to the node phys-schost-1 from the tape device /dev/rmt/0. The metaset command is run from another node in the cluster, phys-schost-2, to remove and later add back node phys-schost-1 to the disk set schost-1. All other commands are run from phys-schost-1 . A new boot block is created on /dev/rdsk/c0t0d0s0, and three state database replicas are re-created on /dev/rdsk/c0t0d0s4 .

[*Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on a cluster node other than the node to be restored* .]
[*Remove the node from the metaset*:]
```
phys-schost-2# metaset -s schost-1 -f -d -h phys-schost-1
```
[*Replace the failed disk and boot the node*:]
Restore the root (/) and /usr *file system using the procedure in the Solaris system administration documentation*
 [*Reboot*:]
```
# reboot
```
[*Replace the disk ID*:]
```
# cldevice repair /dev/dsk/c0t0d0
```
[*Re-create state database replicas*:]
```
# metadb -c 3 -af /dev/rdsk/c0t0d0s4
```
[*Add the node back to the metaset*:]

phys-schost-2# **metaset -s schost-1 -a -h phys-schost-1**

▼ How to Restore a Root (/) File System That Was on a Solstice DiskSuite Metadevice or Solaris Volume Manager Volume

Use this procedure to restore a root (/) file system that was on a Solstice DiskSuite metadevice or a Solaris Volume Manager volume when the backups were performed. Perform this procedure under circumstances such as when a root disk is corrupted and replaced with a new disk. The node being restored should not be booted. Ensure that the cluster is running without errors before performing the restore procedure.

Note – Because you must partition the new disk by using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and re-create file systems as appropriate.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Become superuser or assume a role that provides** solaris.cluster.modifiy **RBAC authorization on a cluster node with access to the disk set,** *other than* **the node you restoring.**

Use a node *other than* the node that you are restoring.

2 **Remove the host name of the node being restored from all disk sets.**

metaset -s *setname* **-f -d -h** *nodelist*

-s *setname*   Specifies the metaset name.

-f   Deletes the last host from the diskset.

-d   Deletes from the metaset.

-h *nodelist*   Specifies the name of the node to delete from the metaset.

3 **Replace the failed disk on the node on which the root (/) file system will be restored.**

Refer to disk replacement procedures in the documentation that shipped with your server.

4 **Boot the node that you are restoring.**

- If you are using the Solaris OS CD, note the following:

  - SPARC: Type:

    ok **boot cdrom -s**

- x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. In the Current Boot Parameters screen, type b or i.

```
                        <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
7,1/sd@0,0:a
Boot args:

Type b [file-name] [boot-flags] <ENTER> to boot with options
or   i <ENTER>                          to enter boot interpreter
or   <ENTER>                            to boot with defaults

                     <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -s
```

- If you are using a Solaris JumpStart™ server, note the following:

  - SPARC: Type:

    ```
    ok boot net -s
    ```

  - x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. In the Current Boot Parameters screen, type b or i.

    ```
                            <<< Current Boot Parameters >>>
    Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
    7,1/sd@0,0:a
    Boot args:

    Type b [file-name] [boot-flags] <ENTER> to boot with options
    or   i <ENTER>                          to enter boot interpreter
    or   <ENTER>                            to boot with defaults

                         <<< timeout in 5 seconds >>>
    Select (b)oot or (i)nterpreter: b -s
    ```

5 Create all the partitions and swap space on the root disk by using the format **command.**

Re-create the original partitioning scheme that was on the failed disk.

6 Create the root (/) file system and other file systems as appropriate, by using the newfs **command**

Re-create the original file systems that were on the failed disk.

Note – Ensure that you create the /global/.devices/node@nodeid file system.

---

7 Mount the root (/) file system on a temporary mount point.

mount *device temp-mountpoint*

8   Use the following commands to restore the root (/) file system.

```
# cd temp-mountpoint
# ufsrestore rvf dump-device
# rm restoresymtable
```

9   Install a new boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/'uname -i'/lib/fs/ufs/bootblk
raw-disk-device
```

10   Remove the lines in the /temp-mountpoint/etc/system **file for MDD root information.**

```
* Begin MDD root info (do not edit)
forceload: misc/md_trans
forceload: misc/md_raid
forceload: misc/md_mirror
forceload: misc/md_hotspares
forceload: misc/md_stripe
forceload: drv/pcipsy
forceload: drv/glm
forceload: drv/sd
rootdev:/pseudo/md@0:0,10,blk
* End MDD root info (do not edit)
```

11   Edit the /temp-mountpoint/etc/vfstab **file to change the root entry from a Solstice DiskSuite metadevice or a Solaris Volume Manager volume to a corresponding normal slice for each file system on the root disk that is part of the metadevice or volume.**

```
Example:
Change from—
/dev/md/dsk/d10   /dev/md/rdsk/d10    /       ufs   1     no      -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  /       ufs   1     no      -
```

12   Unmount the temporary file system, and check the raw disk device.

```
# cd /
# umount temp-mountpoint
# fsck raw-disk-device
```

13   Reboot the node in multiuser mode.

```
# reboot
```

14   Replace the disk ID.

```
# cldevice repair rootdisk
```

15   Use the metadb **command to re-create the state database replicas.**

    # **metadb -c** *copies* **-af** *raw-disk-device*

    -c *copies*   Specifies the number of replicas to create.

    -af *raw-disk-device*   Creates initial state database replicas on the named raw disk device.

16   From a cluster node other than the restored node, add the restored node to all disk sets.

    phys-schost-2# **metaset -s** *setname* **-a -h** *nodelist*

    -a   Adds (creates) the metaset.

    Set up the metadevice or volume/mirror for root (/) according to the Solstice DiskSuite
    documentation.

    The node is rebooted into cluster mode. The cluster is ready to use.

**Example 11–7   Restoring a Root (/) File System That Was on a Solstice DiskSuite Metadevice or
Solaris Volume Manager Volume**

    The following example shows the root (/) file system restored to the node phys-schost-1 from the
    tape device /dev/rmt/0. The metaset command is run from another node in the cluster,
    phys-schost-2, to remove and later add back node phys-schost-1 to the metaset schost-1. All
    other commands are run from phys-schost-1 . A new boot block is created on
    /dev/rdsk/c0t0d0s0, and three state database replicas are re-created on /dev/rdsk/c0t0d0s4 .

[*Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on a cluster node with access to the metaset,
other than the node to be restored*.]
[*Remove the node from the metaset*:]
phys-schost-2# **metaset -s schost-1 -f -d -h phys-schost-1**
[*Replace the failed disk and boot the node*:]

            Boot the node from the Solaris OS CD:

            ■   SPARC: Type:

                ok **boot cdrom -s**

            ■   x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then
                turning it off and on. In the Current Boot Parameters screen, type b or i.

```
                        <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type  b [file-name] [boot-flags] <ENTER> to boot with options
or    i <ENTER>                          to enter boot interpreter
or    <ENTER>                            to boot with defaults
```

```
                              <<< timeout in 5 seconds >>>
    Select (b)oot or (i)nterpreter: b -s
```

[*Use* format *and* newfs *to re-create partitions and file systems*
.]
[*Mount the root file system on a temporary mount point*:]
 # mount /dev/dsk/c0t0d0s0 /a
[*Restore the root file system*:]
 # cd /a
 # ufsrestore rvf /dev/rmt/0
 # rm restoresymtable
[*Install a new boot block*:]
**# /usr/sbin/installboot /usr/platform/'uname **
-i'/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0

[*Remove the lines in / temp-mountpoint/etc/system file for MDD root information:*
]
```
* Begin MDD root info (do not edit)
forceload: misc/md_trans
forceload: misc/md_raid
forceload: misc/md_mirror
forceload: misc/md_hotspares
forceload: misc/md_stripe
forceload: drv/pcipsy
forceload: drv/glm
forceload: drv/sd
rootdev:/pseudo/md@0:0,10,blk
* End MDD root info (do not edit)
```
[*Edit the* /temp-mountpoint/etc/vfstab *file*]
```
Example:
Change from—
/dev/md/dsk/d10   /dev/md/rdsk/d10   /      ufs   1    no       -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0 /usr   ufs   1    no       -
```
[*Unmount the temporary file system and check the raw disk device:*]
 # cd /
 # umount /a
 # fsck /dev/rdsk/c0t0d0s0
[Reboot:]
 # reboot
[Replace the disk ID:]
 # cldevice repair /dev/rdsk/c0t0d0
[Re-create state database replicas:]
 # metadb -c 3 -af /dev/rdsk/c0t0d0s4
[*Add the node back to the metaset*:]

```
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1
```

▼ SPARC: How to Restore a Non-Encapsulated Root (/) File System (VERITAS Volume Manager)

Use this procedure to restore a non-encapsulated root (/) file system to a node. The node being restored should not be booted. Ensure the cluster is running without errors before performing the restore procedure.

Note – Because you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and re-create file systems as appropriate.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1 **Replace the failed disk on the node where the root file system will be restored.**
   Refer to disk replacement procedures in the documentation that shipped with your server.

2 **Boot the node that you are restoring.**
   - If you are using the Solaris OS CD, at the OpenBoot PROM ok prompt, type the following command:

     ok **boot cdrom -s**
   - If you are using a Solaris JumpStart server, at the OpenBoot PROM ok prompt, type the following command:

     ok **boot net -s**

3 **Create all the partitions and swap on the root disk by using the** format **command.**
   Re-create the original partitioning scheme that was on the failed disk.

4 **Create the root (/) file system and other file systems as appropriate, using the** newfs **command.**
   Re-create the original file systems that were on the failed disk.

Note – Ensure that you create the /global/.devices/node@nodeid file system.

---

5 **Mount the root (/) file system on a temporary mount point.**
   # **mount** *device temp-mountpoint*

6 Restore the root (/) file system from backup, and unmount and check the file system.

```
# cd temp-mountpoint
# ufsrestore rvf dump-device
# rm restoresymtable
# cd /
# umount temp-mountpoint
# fsck raw-disk-device
```

The file system is now restored.

7 Install a new boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/'uname -i'/lib/fs/ufs/bootblk raw-disk-device
```

8 Reboot the node in multiuser mode.

```
# reboot
```

9 Update the disk ID.

```
# cldevice repair /dev/rdsk/disk-device
```

10 Press Control-d to resume in multiuser mode.

The node reboots into cluster mode. The cluster is ready to use.

Example 11–8 SPARC: Restoring a Non-Encapsulated Root (/) File System (VERITAS Volume Manager)

The following example shows a non-encapsulated root (/) file system that is restored to the node phys-schost-1 from the tape device /dev/rmt/0.

[*Replace the failed disk and boot the node*:]

Boot the node from the Solaris OS CD. At the OpenBoot PROM ok prompt, type the following command:

```
ok boot cdrom -s
...
```
[*Use* format *and* newfs *to create partitions and file systems*]
[*Mount the root file system on a temporary mount point*:]
```
# mount /dev/dsk/c0t0d0s0 /a
```
[*Restore the root file system*:]
```
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0
```

[*Install a new boot block*:]
```
# /usr/sbin/installboot /usr/platform/`uname \
-i`/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0
```

[*Reboot*:]
```
# reboot
```
[*Update the disk ID*:]
```
# cldevice repair /dev/rdsk/c0t0d0
```

▼ SPARC: How to Restore an Encapsulated Root (/) File System (VERITAS Volume Manager)

Use this procedure to restore an encapsulated root (/) file system to a node. The node being restored should not be booted. Ensure the cluster is running with errors before performing the restore procedure.

Note – Because you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and re-create file systems as appropriate.

This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix A.

1   Replace the failed disk on the node where the root file system will be restored.

Refer to disk replacement procedures in the documentation that shipped with your server.

2   Boot the node that you are restoring.

- If you are using the Solaris OS CD, at the OpenBoot PROM ok prompt, type the following command:

  ```
  ok boot cdrom -s
  ```

- If you are using a Solaris JumpStart server, at the OpenBoot PROM ok prompt, type the following command:

  ```
  ok boot net -s
  ```

3   Create all the partitions and swap space on the root disk by using the format **command.**

Re-create the original partitioning scheme that was on the failed disk.

4   Create the root (/) file system and other file systems as appropriate, by using the newfs **command.**

Re-create the original file systems that were on the failed disk.

Note – Ensure that you create the /global/.devices/ *node@nodeid* file system.

---

5 Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mountpoint
```

6 Restore the root (/) file system from backup.

```
# cd temp-mountpoint
# ufsrestore rvf dump-device
# rm restoresymtable
```

7 Create an empty install-db **file.**

This file puts the node in VxVM installation mode at the next reboot.

```
# touch \
/temp-mountpoint/etc/vx/reconfig.d/state.d/install-db
```

8 Remove the following entries from the / *temp-mountpoint*/etc/system **file.**

```
* rootdev:/pseudo/vxio@0:0
* set vxio:vol_rootdev_is_volume=1
```

9 Edit the / *temp-mountpoint* /etc/vfstab **file and replace all VxVM mount points with the standard disk devices for the root disk, such as** /dev/dsk/c0t0d0s0**.**

```
Example:
Change from—
/dev/vx/dsk/rootdg/rootvol /dev/vx/rdsk/rootdg/rootvol /     ufs  1     no -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  / ufs   1     no       -
```

10 Unmount the temporary file system and check the file system.

```
# cd /
# umount temp-mountpoint
# fsck raw-disk-device
```

11 Install the boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/'uname -i'/lib/fs/ufs/bootblk raw-disk-device
```

12 Reboot the node in multiuser mode.

```
# reboot
```

13 Update the disk ID by using scdidadm **(1M).**

```
# cldevice repair /dev/rdsk/c0t0d0
```

14   Run the vxinstall **command to encapsulate the disk and reboot.**

15   If a conflict in minor number occurs with any other system, unmount the global devices and reminor the disk group.

- Unmount the global devices file system on the cluster node.

    # **umount /global/.devices/**node@nodeid

- Reminor the rootdg disk group on the cluster node.

    # **vxdg reminor rootdg 100**

16   Shut down and reboot the node in cluster mode.

    # **shutdown -g0 -i6 -y**

Example 11–9   SPARC: Restoring an Encapsulated root (/) File System (VERITAS Volume Manager)

The following example shows an encapsulated root (/) file system restored to the node phys-schost-1 from the tape device /dev/rmt/0.

[*Replace the failed disk and boot the node*:]

Boot the node from the Solaris OS CD. At the OpenBoot PROM ok prompt, type the following command:

```
ok boot cdrom -s
...
[Use format and newfs to create partitions and file systems]
[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a
[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
[Create an empty install-db file:]
# touch /a/etc/vx/reconfig.d/state.d/install-db
[Edit /etc/system on the temporary file system and
remove or comment out the following entries:]
    # rootdev:/pseudo/vxio@0:0
    # set vxio:vol_rootdev_is_volume=1
[Edit /etc/vfstab on the temporary file system:]
Example:
Change from—
/dev/vx/dsk/rootdg/rootvol /dev/vx/rdsk/rootdg/rootvol / ufs 1 no-

Change to—
```

```
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  / ufs  1   no      -
```
[*Unmount the temporary file system, then check the file system*:]
cd /
umount /a
fsck /dev/rdsk/c0t0d0s0
[*Install a new boot block*:]
**/usr/sbin/installboot /usr/platform/'uname **
-i'/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0

[*Reboot*:]
reboot
[*Update the disk ID*:]
cldevice repair /dev/rdsk/c0t0d0
[*Encapsulate the disk:*:]
vxinstall
Choose to encapsulate the root disk.
[*If a conflict in minor number occurs, reminor the rootdg disk group
*:]
umount /global/.devices/node@nodeid
vxdg reminor rootdg 100
shutdown -g0 -i6 -y

See Also   For instructions about how to mirror the encapsulated root disk, see the *Sun Cluster Software Installation Guide for Solaris OS*.

12

Administering Sun Cluster With the Graphical User Interfaces

This chapter provides descriptions of Sun Cluster Manager and Sun Management Center graphical user interface (GUI) tools, which you can use to administer many aspects of a cluster. It also contains procedures to configure and start Sun Cluster Manager. The online help that is included with each GUI provides instructions for accomplishing various administrative tasks by using the GUI.

This chapter includes the following:

- "Overview of Sun Cluster Manager" on page 313
- "SPARC: Overview of Sun Management Center" on page 314
- "Configuring Sun Cluster Manager" on page 315
- "Starting the Sun Cluster Manager Software" on page 318

Overview of Sun Cluster Manager

Sun Cluster Manager is a GUI that enables you to graphically display cluster information, monitor configuration changes, and check the status of cluster components. Sun Cluster Manager also enables you to perform many administrative tasks for the following Sun Cluster components.

- Adapters
- Cables
- Data services
- Global devices
- Interconnects
- Junctions
- Nodes
- Quorum devices
- Resource groups
- Resources

Information about installing and using Sun Cluster Manager can be found in the following locations.

- **Installing and starting Sun Cluster Manager:** See the *Sun Cluster Software Installation Guide for Solaris OS*.

- **Configuring port numbers, server addresses, security certificates, and users:** See "Configuring Sun Cluster Manager" on page 315.

- **Installing and administering aspects of your cluster using Sun Cluster Manager:** See the online help supplied with Sun Cluster Manager.

- **Regenerating Sun Cluster Manager security keys:** See "How to Regenerate Common Agent Container Security Keys" on page 316.

Note – However, Sun Cluster Manager currently cannot perform all Sun Cluster administrative tasks. You must use the command-line interface for some operations.

SPARC: Overview of Sun Management Center

The Sun Cluster module for Sun Management Center™ (formerly Sun Enterprise SyMON™) GUI console enables you to graphically display cluster resources, resource types, and resource groups. It also enables you to monitor configuration changes and check the status of cluster components. However, the Sun Cluster module for Sun Management Center cannot perform Sun Cluster configuration tasks. You must use the command-line interface for configuration operations. See "Command-Line Interface" in Chapter 1 for more information.

For information about installing and starting the Sun Cluster module for Sun Management Center, and for viewing the cluster-specific online help supplied with the Sun Cluster module, see the *Sun Cluster Software Installation Guide for Solaris OS*.

The Sun Cluster module of Sun Management Center is Simple Network Management Protocol (SNMP) compliant. Sun Cluster has created a Management Information Base (MIB) that can be used as the data definition by third-party management stations based on SNMP.

The Sun Cluster MIB file is located at `/opt/SUNWsymon/modules/cfg/sun-cluster-mib.mib` on any cluster node.

The Sun Cluster MIB file is an ASN.1 specification of the Sun Cluster data that is modeled. This is the same specification used by all Sun Management Center MIBs. To use the Sun Cluster MIB, refer to the instructions for using other Sun Management Center MIBs in the "SNMP MIBs for Sun Management Center Modules" in *Sun Management Center 3.6 User's Guide* in "SNMP MIBs for Sun Management Center Modules" in *Sun Management Center 3.6 User's Guide*.

Configuring Sun Cluster Manager

Sun Cluster Manager is a GUI that you can use to administer and view the status of all aspects of quorum devices, IPMP groups, interconnect components, and global devices. You can use the GUI in place of many of the Sun Cluster CLI commands.

The procedure for installing Sun Cluster Manager on your cluster is included in the *Sun Cluster Software Installation Guide for Solaris OS*. The Sun Cluster Manager online help contains instructions for completing various tasks by using the GUI.

This section contains the following procedures for reconfiguring Sun Cluster Manager after initial installation.

- "Setting up RBAC Roles" on page 315
- "How to Change the Server Address for Sun Cluster Manager" on page 316
- "How to Regenerate Common Agent Container Security Keys" on page 316

Setting up RBAC Roles

The Sun Cluster Manager uses RBAC to determine who has rights to administer the cluster. Several RBAC rights profiles are included in the Sun Cluster software. You can assign these rights profiles to users or to roles to give users different levels of access to Sun Cluster. For more information about how to set up and manage RBAC for Sun Cluster, see Chapter 2.

▼ How to Use the Common Agent Container to Change the Port Numbers for Services or Management Agents

If the default port numbers for your common agent container services conflict with other running processes, you can use the `cacaoadm` command to change the port number of the conflicting service or management agent on each node of the cluster.

1 On all cluster nodes, stop the common agent container management daemon.

```
# /opt/SUNWcacao/bin/cacaoadm stop
```

2 Stop Sun Java Web Console.

```
# /usr/sbin/sunmcwebserver stop
```

3 Retrieve the port number currently used by the common agent container service with the `get-param` **subcommand.**

```
# /opt/SUNWcacao/bin/cacaoadm get-param parameterName
```

You can use the `cacaoadm` command to change the port numbers for the following common agent container services. The following list provides some examples of services and agents that can be managed by the common agent container, along with corresponding parameter names.

JMX connector port	`jmxmp-connector-port`
SNMP port	`snmp-adaptor-port`
SNMP trap port	`snmp-adaptor-trap-port`
Command stream port	`commandstream-adaptor-port`

4 Change a port number.

```
# /opt/SUNWcacao/bin/cacaoadm set-param parameterName=parameterValue
=parameterValue
```

5 Repeat Step 4 on each node of the cluster.

6 Restart Sun Java Web Console.

```
# /usr/sbin/sunmcwebserver start
```

7 Restart the common agent container management daemon on all cluster nodes.

```
# /opt/SUNWcacao/bin/cacaoadm start
```

▼ How to Change the Server Address for Sun Cluster Manager

If you change the host name of a cluster node, you must change the address from which Sun Cluster Manager runs. The default security certificate is generated based on the node's host name at the time Sun Cluster Manager is installed. To reset the node's host name, delete the certificate file, keystore and restart Sun Cluster Manager. Sun Cluster Manager automatically creates a new certificate file with the new host name. You must complete this procedure on any node that has had its host name changed.

1 Remove the certificate file, keystore, **located in** /etc/opt/webconsole**.**

```
# cd /etc/opt/webconsole
# pkgrm keystore
```

2 Restart Sun Cluster Manager.

```
# /usr/sbin/smcwebserver restart
```

▼ How to Regenerate Common Agent Container Security Keys

Sun Cluster Manager uses strong encryption techniques to ensure secure communication between the Sun Cluster Manager web server and each cluster node.

The keys that Sun Cluster Manager uses are stored under the `/etc/opt/SUNWcacao/security` directory on each node. They should be identical across all cluster nodes.

Under normal operation, these keys can be left in their default configuration. If you change the host name of a cluster node, you must regenerate the common agent container security keys. You might also need to regenerate the keys because a possible key compromise (for example, root compromise on the machine). To regenerate the security keys, use the following procedure.

1   On all cluster nodes, stop the common agent container management daemon.

```
# /opt/SUNWcacao/bin/cacaoadm stop
```

2   On one node of the cluster, regenerate the security keys.

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm create-keys --force
```

3   Restart the common agent container management daemon on the node on which you regenerated the security keys.

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm start
```

4   Create a tar file of the `/etc/cacao/instances/default` **directory.**

```
phys-schost-1# cd /etc/cacao/instances/default
phys-schost-1# tar cf /tmp/SECURITY.tar security
```

5   Copy the `/tmp/Security.tar` **file to each of the cluster nodes.**

6   On each node to which you copied the `/tmp/SECURITY.tar` **file, extract the security files.**

Any security files that already exist in the `/etc/opt/SUNWcacao/` directory are overwritten.

```
phys-schost-2# cd /etc/cacao/instances/default
phys-schost-2# tar xf /tmp/SECURITY.tar
```

7   Delete the `/tmp/SECURITY.tar` **file from each node in the cluster.**

You must delete each copy of the tar file to avoid security risks.

```
phys-schost-1# rm /tmp/SECURITY.tar

phys-schost-2# rm /tmp/SECURITY.tar
```

8   On all nodes, restart the common agent container management daemon.

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm start
```

9   Restart Sun Cluster Manager.

```
# /usr/sbin/smcwebserver restart
```

Starting the Sun Cluster Manager Software

The Sun Cluster Manager graphical user interface (GUI) provides an easy way to administer some aspects of the Sun Cluster software. See the Sun Cluster Manager online help for more information.

▼ How to Start Sun Cluster Manager

This procedure shows how to start Sun Cluster Manager on your cluster.

1. **Do you intend to access Sun Cluster Manager by using the cluster node** `root` **user name and password rather than set up a different user name and password?**

    ■ If yes, go to Step 5.
    ■ If no, go to Step 3 to set up Sun Cluster Manager user accounts.

2. **Become superuser on a cluster node.**

3. **Create a user account to access the cluster through Sun Cluster Manager.**

    You use the `useradd`(1M) command to add a user account to the system. You must set up at least one user account to access Sun Cluster Manager if you do not use the `root` system account. Sun Cluster Manager user accounts are used only by Sun Cluster Manager. These accounts do not correspond to any Solaris OS system user accounts. Creating and assigning an RBAC role to a user account is described in more detail in "Creating and Assigning an RBAC Role With a Sun Cluster Management Rights Profile" on page 45.

    ***

    Note – Users who do not have a user account set up on a particular node cannot access the cluster through Sun Cluster Manager from that node, nor can users manage that node through another cluster node to which the users do have access.

    ***

4. **(Optional) Repeat Step 3 to set up additional user accounts.**

5. **From the administrative console or any other machine outside the cluster, start a browser.**

6. **Ensure that the browser's disk and memory cache sizes are set to a value that is greater than 0.**

7. **Ensure that the Java and Javascript are enables in the browser.**

8. **Accept any certificates that are presented by the web browser.**

    The Java Web Console login page is displayed.

9. **Enter the user name and password for the user you want to access Sun Cluster Manager.**

10 **Click the** Log In **button.**

The Java Web Console application launch page is displayed.

11 **Click the** Sun Cluster Manager **link under the** Systems **category.**

12 **Accept any additional certificates that are presented by the web browser.**

13 **From the browser, connect to the Sun Cluster Manager port on one node of the cluster.**

The default port number is 6789.

https://_node_**:6789/**

A
APPENDIX A

Sun Cluster Object-Oriented Commands

This appendix introduces the object-oriented commands, their short forms, and their subcommands.

Object-Oriented Command Names and Aliases

In addition to their longer and more descriptive forms, many Sun Cluster commands also have a short form, or alias, that significantly reduces the amount you must type. The following table lists the commands and their shorter aliases.

TABLE A–1 Object-Oriented Commands and Aliases (Short Names)

Full Command	Alias	Purpose
claccess	none	Manage Sun Cluster access policies
cldevice	cldev	Manage Sun Cluster devices
cldevicegroup	cldg	Manage Sun Cluster device groups
clinterconnect	clintr	Manage the Sun Cluster interconnect
clnasdevice	clnas	Manage access to NAS devices for Sun Cluster
clnode	none	Manage Sun Cluster nodes
clquorum	clq	Manage Sun Cluster quorum
clquorumserver	clqs	Configure and manage quorum server processes on the quorum server host
clreslogicalhostname	clrslh	Manage Sun Cluster resources for logical host names
clresource	clrs	Manage resources for Sun Cluster data services
clresourcegroup	clrg	Manage resource groups for Sun Cluster data services

TABLE A–1 Object-Oriented Commands and Aliases (Short Names)   *(Continued)*

Full Command	Alias	Purpose
clresourcetype	clrt	Manage resource types for Sun Cluster data services
clrssharedaddress	clrssa	Manage Sun Cluster resources for shared addresses
clsetup	none	Configure Sun Cluster interactively. This command has no subcommands.
clsnmphost	none	Administer Sun Cluster SNMP hosts
clsnmpmib	none	Administer the Sun Cluster SNMP MIB
clsnmpuser	none	Administer Sun Cluster SNMP users
cltelemetryattribute	clta	Configure system resource monitoring.
cluster	none	Manage the global configuration and status of Sun Cluster
clvxvm	none	Configure Veritas Volume Manager for Sun Cluster

Object-Oriented Command Set Overview

The following tables list the commands in the object-oriented command set and the subcommands available with each command.

TABLE A–2 claccess: Manage Sun Cluster Access Policies for Nodes

Subcommand	Purpose
allow	Allows the specified machine or machines access to the cluster configuration.
allow-all	Allows all nodes access to the cluster configuration.
deny	Denies the specified machine or machines access to the cluster configuration.
deny-all	Denies all nodes access to the cluster configuration.
list	Displays the names of the machines that have access to the cluster configuration.
set	Sets the authentication protocol to the value that you specify with the -a option.
show	Displays the names of the machines that have access to the cluster configuration.

TABLE A–3 cldevice, cldev: Manage Sun Cluster Devices

Subcommand	Purpose
check	Performs a consistency check to compare the kernel representation of the devices against the physical devices.

TABLE A–3 `cldevice`, `cldev`: Manage Sun Cluster Devices   *(Continued)*

Subcommand	Purpose
clear	Removes all DID references to underlying devices that are detached from the current node.
combine	Combines the specified DID instance with a new destination instance.
export	Exports configuration information for a cluster device.
list	Displays all device paths.
monitor	Turns on monitoring for the specified disk paths.
populate	Populates the global-devices namespace.
refresh	Updates the device configuration information that is based on the current device trees on a cluster node.
rename	Moves the specified DID instance to a new DID instance.
repair	Performs a repair procedure on the specified device instances.
replicate	Configures DID devices for use with controller-based replication.
set	Sets the properties of the specified device.
show	Displays a configuration report for all specified device paths.
status	Displays the status of the disk paths that are specified as operands to the command.
unmonitor	Turns off monitoring for the disk paths that are specified as operands to the command.

TABLE A–4 `cldevicegroup`, `cldg`: Manage Sun Cluster Device Groups

Subcommand	Purpose
add-device	Adds new member disk devices to an existing raw-disk device group.
add-node	Adds new nodes to an existing device group.
create	Creates a new device group.
delete	Deletes device groups.
disable	Disables offline device groups.
enable	Enables device groups.
export	Exports the device-group configuration information.
list	Displays a list of device groups.
offline	Takes device groups offline.
online	Brings device groups online on a pre-designated node.

TABLE A–4 cldevicegroup, cldg: Manage Sun Cluster Device Groups *(Continued)*

Subcommand	Purpose
remove-device	Removes member disk devices from a raw-disk device group.
remove-node	Removes nodes from existing device groups.
set	Sets attributes that are associated with a device group.
show	Generates a configuration report for device groups.
status	Generates a status report for device groups.
switch	Transfers device groups from one primary node in a Sun Cluster configuration to another node.
sync	Synchronizes device-group information with the clustering software.

TABLE A–5 clinterconnect, clintr: Manage the Sun Cluster Interconnect

Subcommand	Purpose
add	Adds the new cluster interconnect components that are specified as operands to the command.
disable	Disables the interconnect components that are specified as operands to the command.
enable	Enables the interconnect components that are specified as operands to the command.
export	Exports the cluster interconnect configuration information.
remove	Removes the cluster interconnect components that are supplied as operands to the command.
show	Displays the configuration of interconnect components.
status	Displays the status of the interconnect paths.

TABLE A–6 clnasdevice, clnas: Manage Access to NAS Devices for Sun Cluster

Subcommand	Purpose
add	Adds a NAS device to the Sun Cluster configuration.
add-dir	Adds the specified directories of an already configured NAS device to the cluster configuration.
export	Exports the cluster NAS device configuration information.
list	Displays the NAS devices configured in the cluster.
remove	Removes the specified NAS device or devices from the Sun Cluster configuration.
remove-dir	Removes the specified NAS directory or directories from the Sun Cluster configuration.

TABLE A–6 clnasdevice, clnas: Manage Access to NAS Devices for Sun Cluster   *(Continued)*

Subcommand	Purpose
set	Sets specified properties of a specific NAS device.
show	Displays configuration information for NAS devices in the cluster.

TABLE A–7 clnode: Manage Sun Cluster Nodes

Subcommand	Purpose
add	Configures and adds a node to the cluster.
add-farm	Adds a farm node to a cluster.
clear	Removes a node from the Sun Cluster software configuration.
evacuate	Attempts to switch over all resource groups and device groups from the specified node to a new set of primary nodes.
export	Exports the node or farm configuration information to a file or to the standard output (stdout).
list	Displays the names of nodes that are configured in the cluster or in the farm.
remove	Removes a node from the cluster.
remove-farm	Removes a farm node from a cluster.
set	Sets the properties that are associated with the node that you specify.
show	Displays the configuration of the specified node or nodes.
show-rev	Displays the names of and release information about the Sun Cluster packages that are installed on a node.
status	Displays the status of the node or nodes that you specify.

TABLE A–8 clquorum, clq: Manage Sun Cluster Quorum Configuration

Subcommand	Purpose
add	Adds the specified shared device as a quorum device.
disable	Puts a quorum device or node in the quorum maintenance state.
enable	Removes a quorum device or a node from the quorum maintenance state.
export	Exports the configuration information for the cluster quorum.
list	Displays the names of quorum devices that are configured in the cluster.
remove	Removes the specified quorum device or devices from the Sun Cluster quorum configuration.

TABLE A–8 `clquorum`, `clq`: Manage Sun Cluster Quorum Configuration   *(Continued)*

Subcommand	Purpose
reset	Resets the entire quorum configuration to the default vote count settings.
show	Displays the properties of quorum devices.
status	Displays the status and vote counts of quorum devices.

TABLE A–9 `clquorumserver`, `clqs`: Manage Quorum Servers

Subcommand	Purpose
clear	**Removes outdated cluster information from the quorum server.**
show	Displays the configuration information about the quorum server.
start	Starts the quorum server process on the host machine.
stop	Stops the quorum server process.

TABLE A–10 `clreslogicalhostname`, `clrslh`: Manage Resources for Sun Cluster Logical Host Names

Subcommand	Purpose
create	Creates new logical host name resources.
delete	Deletes logical host name resources.
disable	Disables logical host name resources.
enable	Enables logical host name resources.
export	Exports logical host name resource configuration,
list	Displays a list of the logical host name resources.
list-props	Displays a list of the properties of the logical host name resources.
monitor	Turns on monitoring for logical host name resources.
reset	Clears an error flag that is associated with logical host name resources.
set	Sets specified properties of the logical host name resources.
show	Displays the configuration of logical host name resources.
status	Displays the status of logical host name resources.
unmonitor	Turns off monitoring for logical host name resources.

TABLE A–11 `clresource`, `clrs`: Manage Resources for Sun Cluster Data Services

Subcommand	Purpose
create	Creates the resources that are specified as operands to the command.
delete	Deletes the resources that are specified as operands to the command.
disable	Disables resources.
enable	Enables resources.
export	Exports the cluster resource configuration.
list	Displays a list of cluster resources.
list-props	Displays a list of resource properties.
monitor	Turns on monitoring for resources.
reset	Clears error flags that are associated with cluster resources.
set	Sets resource properties.
show	Displays resource configuration.
status	Displays resource status.
unmonitor	Turns off resource monitoring.

TABLE A–12 `clresourcegroup`, `clrg`: Manage Resource Groups for Sun Cluster Data Services

Subcommand	Purpose
add-node	Adds a node to the end of the `Nodelist` property for a resource group.
create	Creates a new resource group.
delete	Deletes a resource group.
evacuate	Brings offline all resource groups on the nodes that you specify with the `-n` option.
export	Writes the configuration information for a resource group to a file or to the standard output (stdout).
list	Displays a list of resource groups.
manage	Brings a resource group that you specify to a managed state.
offline	Brings a resource group that you specify to an offline state.
online	Brings a resource group that you specify to an online state.
quiesce	Brings the specified resource group to a quiescent state.
remaster	Switches a resource group that you specify to its most preferred node.

TABLE A–12 `clresourcegroup`, `clrg`: Manage Resource Groups for Sun Cluster Data Services *(Continued)*

Subcommand	Purpose
remove-node	Removes a node from the `Nodelist` property of a resource group.
restart	Takes a resource group offline and then back online on the same set of primary nodes that originally hosted the resource group.
resume	Clears the suspended state of any suspended resource groups that you specify.
set	Sets the properties that are associated with the resource groups that you specify.
show	Generates a configuration report for resource groups that you specify.
status	Generates a status report for resource groups that you specify.
suspend	Suspends Resource Group Manager (RGM) control over all applications that are managed by a resource group that you specify.
switch	Changes the node, or set of nodes, that is mastering a resource group that you specify.
unmanage	Brings a resource group that you specify to an unmanaged state.

TABLE A–13 `clresourcetype`, `clrt`: Manage Resource Types for Sun Cluster Data Services

Subcommand	Purpose
add-node	Adds the specified nodes to the node list for resource types.
export	Exports the cluster resource-type configuration.
list	Displays a list of resource types.
list-props	Displays a list of the resource extension properties or resource type properties of resource types.
register	Registers resource types.
remove-node	Removes a node from the list of nodes for which the resource types in the operand list are registered.
set	Sets properties of resource types.
show	Displays configuration information about resource types that are registered in the cluster.
unregister	Unregisters resource types.

TABLE A–14 `clressharedaddress`, `clrssa`: Manage Sun Cluster Resources for Shared Addresses

Subcommand	Purpose
create	Creates shared address resources.
delete	Deletes shared address resources.

TABLE A–14 `clresharedaddress`, `clrssa`: Manage Sun Cluster Resources for Shared Addresses *(Continued)*

Subcommand	Purpose
disable	Disables shared address resources.
enable	Enables shared address resources.
export	Exports shared address resource configuration.
list	Displays a list of shared address resources.
list-props	Displays a list of properties of shared address resources.
monitor	Turns on monitoring for shared address resources.
reset	Clears an error flag that is associated with shared address resources.
set	Sets specified properties of shared address resources.
show	Displays the configuration of shared address resources.
status	Displays the status of shared address resources.
unmonitor	Turns off monitoring for shared address resources.

TABLE A–15 `clsnmphost`: Administer the List of Sun Cluster SNMP Hosts

Subcommand	Purpose
add	Adds an SNMP host to the specified node configuration.
export	Exports the SNMP host information from the specified node.
list	Lists the SNMP hosts that are configured on the specified node.
remove	Removes an SNMP host from the node configuration.
show	Displays the SNMP host configuration information on the specified node.

TABLE A–16 `clsnmpmib`: Administer Sun Cluster SNMP MIB

Subcommand	Purpose
disable	Disables one or more of the cluster MIBs on the specified nodes.
enable	Enables one or more cluster MIBs on the specified node.
export	Exports the cluster MIB configuration information.
list	Displays a list of cluster MIBs on the specified nodes.
set	Sets the SNMP protocol setting that is used on one or more of the MIBs.
show	Displays configuration information for MIBs on the specified nodes.

TABLE A–17 `clsnmpuser`: Administer Sun Cluster SNMP Users

Subcommand	Purpose
create	Adds a user to the SNMP user configuration on the specified node.
delete	Deletes an SNMPv3 user from the specified node.
export	Exports the SNMP user information from the specified node.
list	Prints a list of SNMPv3 users that are configured on the specified node.
set	Sets the configuration of a user on the specified node.
set-default	Sets the default user and security level to use when sending traps using SNMPv3.
show	Displays the information about the users on the specified node.

TABLE A–18 `cltelemetryattribute`, `clta`: Configure System Resource Monitoring

Subcommand	Purpose
disable	Disables the specified telemetry attribute for the specified object type.
enable	Enables data collection for the specified telemetry attribute for the specified object types.
export	Exports the configuration of the telemetry attributes of object types and object instances to a file or to the standard output (`stdout`).
list	Displays the telemetry attributes that are configured for the specified object types.
print	Displays system resource usage for the specified telemetry attributes that are enabled for the specified object instances or object types.
set-threshold	Modifies the settings of a threshold for a specified telemetry attribute on a specified object on a node.
show	Displays the properties that are configured for telemetry attributes on object types or object instances.

TABLE A–19 `cluster`: Manage the Global Configuration and Status of a Cluster

Subcommand	Purpose
create	Creates a cluster by using configuration information that is stored in a `clconfigfile` file.
export	Exports the configuration information in a cluster configuration file.
list	Displays the name of the cluster on which you issue the cluster command.
list-cmds	Prints a list of all available Sun Cluster commands.
rename	Renames the cluster on which you issue the cluster command.

TABLE A–19 `cluster:` Manage the Global Configuration and Status of a Cluster   *(Continued)*

Subcommand	Purpose
restore-netprops	Repairs the cluster private-network settings of the cluster on which you issue the cluster command.
set	Sets the properties of the cluster on which you issue the cluster command.
set-netprops	Sets the properties of the cluster private network address.
show	Displays detailed configuration information about cluster components for the specified clusters.
show-netprops	Displays the private network address settings.
shutdown	Shuts down the cluster on which you issue the cluster command in an orderly fashion.
status	Displays the status of cluster components in the specified cluster.

TABLE A–20 `clvxvm:` Configure VERITAS Volume Manager for Sun Cluster

Subcommand	Purpose
encapsulate	Encapsulates the root disk and performs other Sun Cluster-specific tasks.
initialize	Initializes VxVM and performs other Sun Cluster-specific tasks.

Index

dedicated processor set, configuring, 271
device group properties, changing, 161
device groups
    adding, 139
    administration overview, 135
    configuring for data replication, 91
    listing configuration, 165
    maintenance state, 167
    new minor number assignment, 148
    primary ownership, 161
    removing and unregistering, 140, 155
    verifying
        registration, 153
device groups, adding, 138
devices, global, 119-185
disabling SNMP event MIB, 260
disabling transport cables, 218
disaster tolerance, definition, 79
disk groups
    creating, 144-145
    modifying, 147
    registering, 149
    registering configuration changes, 151
disk-path monitoring, 119-185
disk path monitoring, print faulted disk paths, 182
disk paths
    monitoring, 180-181
    unmonitoring, 181-182
displaying configured resources, 24
Domain Name System (DNS)
    guidelines for updating, 86
    updating in data replication, 117
DR, dynamic reconfiguration, 189
dynamic reconfiguration, 120-121
    cluster interconnects, 210
    public network interfaces, 225
    quorum devices, 189

E
enabling SNMP event MIB, 260
enabling transport cables, 217
encapsulating disks, 145
error messages, /var/adm/messages file, 71

/etc/nsswitch.conf file, non-global zone
    modifications, 246
/etc/vfstab file, 39
event MIB
    disabling SNMP (Simple Network Management
        Protocol), 260
    enabling SNMP, 260
event MIB (Management Information Base) protocol,
    changing SNMP (Simple Network Management
    Protocol), 259
example configurations (campus clustering), two-room,
    storage-based replication, 75-78
extension properties for data replication
    application resource, 100, 102
    replication resource, 97, 98

F
failback property, 161
failover applications for data replication
    affinity switchover, 83
    guidelines
        managing failover, 86
        resource groups, 83-85
failover for data replication, managing, 115-117
fair share scheduler
    configuring, 265
    CPU shares configuration, 264
file systems
    backing up, 288
    finding names, 288
    restoring encapsulated root, 308
    restoring non-encapsulated root, 306
    restoring root, 299
        from metadevice, 301
        from volume, 301
file systems for NFS application, configuring for data
    replication, 93-95
files
    /etc/vfstab, 39
    md.conf, 138
    md.tab, 19
    ntp.conf.cluster, 235
    restoring interactively, 298

V